Copyright
by
Luis Ricardo Suazo
2018

The Dissertation Committee for Luis Ricardo Suazo certifies that this is the approved version of the following dissertation:

Classical and Semi-classical Analysis of Klein-Gordon Field in a Signature Transition Spacetime

Committee:

Richard Matzner, Supervisor

Austin Gleeson

Andrew Neitzke

Sonia Paban

Lorenzo Sadun

Classical and Semi-classical Analysis of Klein-Gordon Field in a Signature Transition Spacetime

by

Luis Ricardo Suazo

Dissertation
Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the degree of
Doctor of Philosophy

The University of Texas at Austin
May, 2018

Classical and Semi-classical Analysis of Klein-Gordon Field in a Signature Transition Spacetime

Luis Ricardo Suazo, Ph.D.
The University of Texas at Austin, 2018

Supervisor: Richard Matzner

Classical and semiclassical analyses of a Klein-Gordon field minimally coupled to a fixed, signature-changing (Riemannian to Lorentzian) background are performed. The background spacetime is a toy model used to represent a possible history considered in the Hartle-Hawking no boundary proposal. An extensive study is given of matching techniques to allow continuation of the Klein-Gordon field through the change of signature. The classical analysis focuses on exploring the consequences of imposing regularity conditions on the field in the Riemannian sector. The semiclassical analysis investigates the particle production due to gravitational interaction, assuming that when the universe tunnels into existence, the Klein-Gordon field is in the minimum energy vacuum state. It is found that the particle production rate is nonzero at large wavenumber k, and that the rate of production decreases as k^{-2}.

Contents

List of Figures

Conventions

Notational simplifications

- $\frac{\partial}{\partial x^\mu}$ is often written as ∂_μ (whenever coordinates are obvious)
- ∇_{∂_μ} is often written as ∇_μ

The conventions used in these thesis are as follows:

- signature: (-+++)
- connection coefficients: $\nabla_{\partial_\mu}\partial_\nu = A^\rho{}_{\mu\nu}\partial_\rho$
- Torsion operator: $T(u,v) = \nabla_u v - \nabla_v u - [u,v]$
- Christoffel Symbols: $\Gamma^\rho{}_{\mu\nu} = \frac{1}{2}g^{\rho\alpha}\left(\partial_\mu g_{\nu\alpha} + \partial_\nu g_{\mu\alpha} - \partial_\alpha g_{\mu\nu}\right)$
- Riemann curvature operator: $R(u,v)w = \left(\nabla_u\nabla_v - \nabla_v\nabla_u - \nabla_{[u,v]}\right)w$
- Riemann curvature tensor: $R(\partial_\mu,\partial_\nu)\partial_\beta = R^\alpha{}_{\beta\mu\nu}\partial_\alpha$
- Riemann curvature coefficients: $R^\alpha{}_{\beta\mu\nu} = \partial_\mu\Gamma^\alpha{}_{\nu\beta} - \partial_\nu\Gamma^\alpha{}_{\mu\beta} + \Gamma^\alpha{}_{\mu\rho}\Gamma^\rho{}_{\nu\beta} - \Gamma^\alpha{}_{\nu\rho}\Gamma^\rho{}_{\mu\beta}$
- Ricci tensor: $R_{\beta\nu} = R^\alpha{}_{\beta\alpha\nu}$
- Ricci Scalar: $R = g^{\beta\nu}R_{\beta\nu}$
- Einstein tensor: $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}$

The sign placements agree with MTW. Index positioning agrees with MTW, up to permutations of symmetric indices.

Chapter One: Roadmap

In this thesis, I investigate the behavior of a minimally coupled Klein-Gordon field on a fixed, toy model background universe whose metric transitions from being Riemmanian to Lorentzian. Chapter 2 presents a quick historical motivation for the problem being considered and a basic description of the the background geometry. Chapter 3 then discusses the relationship between wave-mechanics and particle-mechanics via the geometric optics approximation, which serves both as further motivation for the work and a introduction to global asymptotic approximations, which are a mainstay of the thesis. Then the work begins in earnest in chapter 4.

The first part is entirely classical in nature. I focus on understanding how imposing simple regularity conditions on the Klein-Gordon field in the Riemannian part of the universe affects its evolution in the Lorentzian part. To this end I take special care in understanding the behavior of the field as it crosses the threshold between the domains of different signature. The metric becomes degenerate at this threshold and the relevant equations become singular. In the simple case of a two dimensional spacetime and a massless field, the Klein-Gordon equation, on the toy-model spacetime, can be solved exactly. Due to the singular nature of the transition, simple matching of boundary conditions cannot be performed at this interface. However, by using a local Frobenius series approximation about the transition surface the solutions in the Riemannian domain can be matched to solutions in the Lorentzian domain. This already leads to an interesting conclusion about the nature of the problem: imposing regularity of the field in the Riemannian domain implies that the field must be complex valued in the Lorentzian domain. Moreover, if the complex nature of the problem is accepted at the outset, then the matching problem disappears altogether, and the singular transition boundary can effectively be treated as a regular point in the domain. This will be discussed in much greater detail in chapter 6.

Things become more complicated for general dimensions and mass, where analytic solutions to the Klein-Gordon equation of motion cannot be obtained. To accomplish the matching in this situation, a slew of local, global, and intermediate asymptotic approximations are patched together. This requires taking many different limits and expansions of many solutions to simple differential equations. As a result it can be

easy to lose track of the the overall picture, and what the relevance of the immediate step might be. For this reason I provide roadmap, summarizing the main domains and expansions considered:

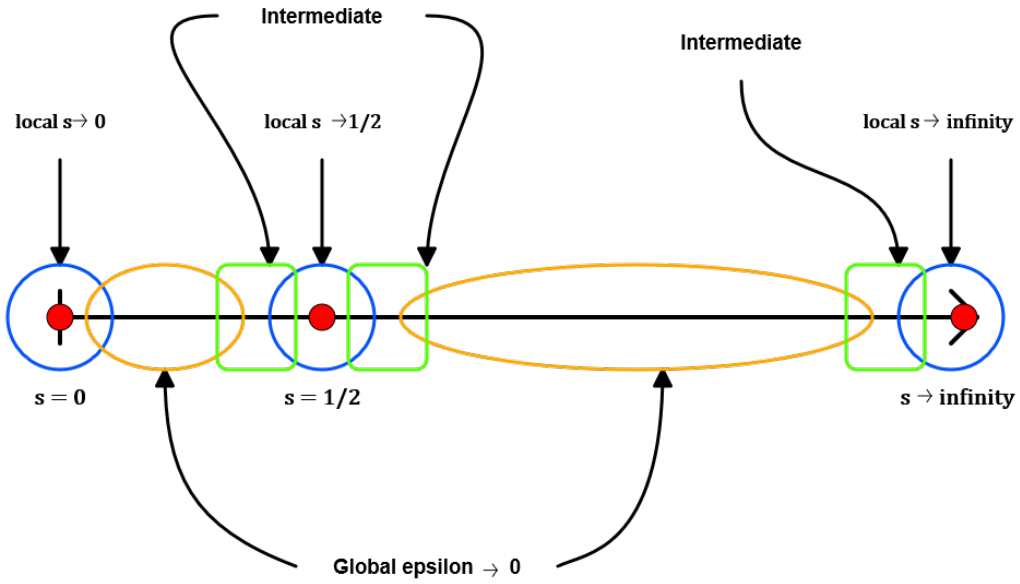
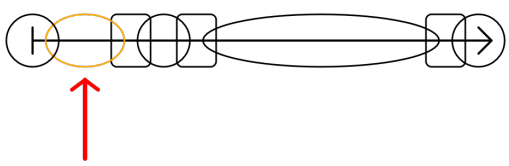
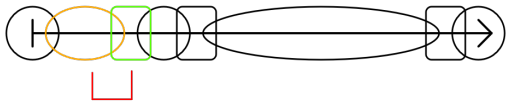


Figure 1.1: These are the separate domains where approximate solutions are obtained throughout the thesis.

The meaning and context of figure (1.1) will become clear as I present each step in the process. However, displaying it ahead of time will make it easy to reference at any point in the discussion. Furthermore, as another perhaps more useful aid, I place smaller, simplified versions of this roadmap (and a small variation of it) throughout the thesis. These serve as signposts at the beginning of sections where expansions are being performed, or a match is attempted, highlighting the particular pieces under consideration. For example, the signpost



means that an expansion about a single domain is being developed. That particular domain is highlighted by both the color of the oval (every other domain representative is black), and the red arrow pointing to the same oval. On the other hand the signpost



is meant to convey that two different expansions, in two different domains are being matched at their overlap.

After considerably more work than in the simplified 2D massless scenario, I show that the situation for the more general model is largely the same. The matching is independent of the dimensions and mass. However, this setting allows for a clear understanding of how the geometric optics approximation fails near the transition boundary. I take this to hint that a more appropriate setting for this analysis should be quantum mechanical in nature. This leads me into the second part of the thesis, which is a semiclassical exploration of particle production in the context of quantum field theories in curved spacetime.

In this second part, I take one step towards a simple semiclassical understanding of scalar fields in signature-transition universes. The question I try to solve is the following: Assuming that when the universe tunnels into existence, the Klein-Gordon field is in its vacuum state (to be made precise in ch 11), what is the particle production due to gravitational effects, as seen at late times? Due to the lack of analytic solutions and the singular nature of the boundaries, answering this question is not so straightforward. To obtain an estimate for the particle production, I extend the work already done in the classical case and use a global geometric optics approximation, along with several other local, and intermediate approximations to bridge solutions at the transition surface with solutions as time goes to infinity. At first glance, this approach seems like it doesn't have much of a chance. The geometrics optics approximation is a limit in which the wave number of the solution goes to infinity. In

this regime, curvature becomes negligible and hence its effect on particle production is negligible. However, as it turns out, the method is able to capture the fact that there is particle production, and that its density goes like one over the inverse wave number squared.

Chapter Two: Setup

2.1 Context: Hartle-Hawking No-Boundary Proposal

In 1973, soon after the discovery of the cosmic microwave background put the big bang model of cosmology on solid footing, Edward Tryon put forth the theory that our universe might have arisen as a vacuum fluctuation (Tryon [1973]). Based on semi-Newtonian arguments about discrete and continuous conservation laws he argued that his proposal roughly predicts a universe much like our own. This idea lay mostly dormant for a few years. Then, after the discovery that Euclidean path integrals and instantons can be used to compute tunneling probabilities (Coleman [1979]) and can be applied to understanding false vacuum decays, particularly in the presence of gravitational forces (Coleman and De Luccia [1980]), Tryon's vacuum fluctuation hypothesis was picked up again. Almost simultaneously, papers by Atkatz and Pagels (Atkatz and Pagels [1982]), Hawking and Moss (Hawking and Moss [1982]) and Vilenkin (Vilenkin [1982]) picked up the thread of universe-wide quantum tunneling. One particular motivation for this work was the appealing possibility of non-singular cosmogenesis; that is, it represented a way to avoid the fate of the Hawking-Penrose singularity theorems (Hawking and Penrose [1970]), as they applied to the big-bang singularity. Then, further inspired by the fact that the Euclidean path integrals provide an alternate[1] way to compute ground state wave functions of a quantum mechanical system (Rossi and Testa [1983]), Hartle and Hawking put forth the no-boundary wave function of the universe (Hartle and Hawking [1983]).

In their seminal paper, Hartle and Hawking propose using the Euclidean path integral to define the ground state wave function of the universe. They argue that since in a quantum theory of a closed universe there is no natural definition of energy, this means that the ground state cannot be defined as the state of lowest energy. In the sum over histories approach to quantum gravity, the wave function is really a

[1] That is, alternate to the canonical Hamiltonian formalism in which a minimal energy eigenstate is sought.

functional over three-metrics and matter-field configurations on a spatial slice (the universe at one moment in time). The path integral can be used to compute probability amplitudes for transitions between an initial three-metric and matter field configuration and a final three-metric and matter field configuration. The integral is taken over all four-geometries that match the initial and final configuration at the boundaries. While the path integral provides the dynamics of the theory, the specification of the initial configuration is tantamount to choosing the actual state of the system, which is necessary for the theory to have any predictive power[2]. The question Hartle and Hawking address is, what is the choice of initial configuration that corresponds to the ground-state wavefunction of the universe? What they propose is that there should be no initial boundary. That is, the path integral sum for the ground-state wavefunction should be carried over compact four geometries that have the final metric and matter configuration as its only boundary. Moreover, Hartle and Hawking state that "one can interpret the functional integral over all compact four-geometries bounded by a given three-geometry as giving the amplitude for that three-geometry to arise from a zero three-geometry, i.e., a single point."

After laying some ground work, Hartle and Hawking implement this proposal in a minusuperspace model with homogeneous, isotropic, and closed S^3 spatial slices. They then use a semiclassical approximation to evaluate the ground-state wavefunction. In particular, they use the fact that in the semiclassical limit, the path integral is dominated by the saddle points of the (Euclidean) action. This way the no-boundary proposal gets translated into a "no-boundary" condition for spacetimes. This leads to considerations of signature changing spacetimes where Lorentzian spacetimes, at early times, give way to Riemannian manifolds without boundaries or singularities.

This Hartle-Hawking no-boundary proposal has served as motivation for the investigation detailed in this thesis. I focus not on a sum over histories approach, but rather on a single signature changing spacetime; a toy model chosen for its simplicity. I consider a Klein-Gordon field on such a spacetime and the effects that a Hartle-Hawking no boundary condition has on its evolution, with particular focus on its behavior across the interface between the Riemannian and Lorentzian sectors. I consider the behavior of said Klein-Gordon field both classically and semi-classically.

Similar signature changing spacetimes have come up in other contexts, namely in

[2]See Halliwell [1991] for a great discussion about this topic.

considerations of topology changes in space, like the trousers topology (see Anderson and DeWitt [1986], Manogue et al. [1988], and Dowker [2003]). The problem of setting up a scalar field and the calculation of particle production in such spaces has also been considered (Dray et al. [1991], Dray et al. [1995], and Dray et al. [1997]). The vast differences between the work detailed in this thesis and similar work in the literature comes from the use of a concrete toy model, instead of general considerations of the theory.

2.2 2D Model Spacetime

The toy model spacetime considered in this chapter (and the following ones) is defined as a submanifold of $\mathbb{R}^n$ with its metric induced from the Minkowski metric on the ambient space. For the two-dimensional model in particular, the starting point is $\mathbb{R}^3$ with a Minkowski metric, which expressed in cylindrical polar coordinates (T, ρ, θ) takes the form

$$-dT \otimes dT + d\rho \otimes d\rho + \rho^2 d\theta \otimes d\theta.$$

The model spacetime under consideration is the hypersurface (paraboloid) defined by the equation $T = \rho^2$. Using the coordinates (ρ, θ) of the ambient space as coordinates for the paraboloid, the induced metric takes the form

$$g = \left(1 - 4\rho^2\right) d\rho \otimes d\rho + \rho^2 d\theta \otimes d\theta.$$

The cylindrical polar coordinates (T, ρ, θ) do not cover $\mathbb{R}^3$ completely. In particular the points approached as $\rho = 0$ are not covered. This behavior is inherited by the (ρ, θ) coordinates on the paraboloid. So the singular behavior at $\rho = 0$ of the induced metric (and the Minkowski metric on $\mathbb{R}^3$) is completely expected, and merely an artifice of the coordinate choice. On the other hand, the points with coordinate value $\rho = 1/2$ are points at which the coordinate system is perfectly valid, except those at the edges of the θ coordinate. So the singular behavior at $\rho = 1/2$ is inherently a metric behavior.

Note that for $\rho < 1/2$, the metric is Riemannian and ρ is a space-like coordinate, while for $\rho > 1/2$ the metric is Lorentzian and ρ is a time-like coordinate. At $\rho = 1/2$ the metric is degenerate. The Riemannian and Lorentzian domains are

defined (respectively) as follows:

$$\begin{aligned} D_R &:= \left\{ \rho \mid 0 \leq \rho \leq \frac{1}{2} \right\} \\ D_L &:= \left\{ \rho \mid \frac{1}{2} \leq \rho < \infty \right\}. \end{aligned}$$

For convenience, the transition point $\rho = 1/2$ is included in both domains. Since most of the analysis performed on this model will take place on the Lorentzian domain D_L, where ρ is a time-like coordinate, the coordinate label t will be used instead of ρ. Therefore the metric will be expressed as

$$g = -\left(4t^2 - 1\right) dt \otimes dt + t^2 d\theta \otimes d\theta.$$

While t is a time-like coordinate in the Lorentzian domain, it is not *cosmic time*. That means, among other things, that given a particle on a path with fixed θ coordinate, t will not measure the proper time elapsed along this curve. It is often preferable to work with such a cosmic time coordinate due to its more immediate physical meaning. In terms of t, the cosmic time coordinate is defined by the condition

$$d\tau = \sqrt{4t^2 - 1} dt. \tag{2.1}$$

Using the coordinate τ the metric takes the form

$$g = -d\tau \otimes d\tau + t(\tau)^2 d\theta \otimes d\theta.$$

However, while the condition relating t and τ can be integrated to obtain (ignoring the constant of integration)

$$\tau = \frac{1}{2} t\sqrt{4t^2 - 1} - \frac{1}{4} \ln\left(2t + \sqrt{4t^2 - 1}\right),$$

it cannot be solved for t in terms of τ. This means that the metric cannot be expressed in closed form, using the τ coordinate. Thus to move forward I continue to use the (t, θ) coordinate system. Note that the derivative of cosmic time with respect to t is

$$\frac{d\tau}{dt} = \sqrt{4t^2 - 1},$$

which goes to zero as t approaches $1/2$. This means that (proper) time "slows down" the closer and closer t gets to $1/2$, for particles on constant θ curves. Close to the

transition, a finite amount of t-time becomes and infinitesimally small amount of τ-time. On the other hand, as t gets very large, a small amount of t-time becomes a large amount of τ time.

This behavior corresponds to a closing off of the light cones. That is, consider the null paths at a given point in spacetime. From the metric, it follows that they satisfy

$$\frac{d\theta}{dt} = \pm\frac{\sqrt{4t^2-1}}{t}. \tag{2.2}$$

As t approaches $1/2$ from above, these velocities tend to zero, and at $t = 1/2$ they coincide. This signifies that the light cones have squeezed shut, which in particular means that there are no time-like path coming out of the singular surface; there are only space-like paths and a single null path (up to reparameterizations). This already portends an ill-fate for any attempt to understand the behavior of matter across the transition surface.

The typical general relativity tensors, the Levi-Civita connection[3], Riemann curvature tensor, Ricci tensor, and the Ricci scalar are:

- Connection coefficients

$$\Gamma^0{}_{00} = \frac{4t}{4t^2-1} \qquad \Gamma^0{}_{11} = \frac{t}{4t^2-1} \qquad \Gamma^1{}_{10} = \Gamma^1{}_{01} = \frac{1}{t}$$

- Riemann tensor

$$R^0{}_{110} = -R^0{}_{101} = \frac{4t^2}{(4t^2-1)^2} \qquad R^1{}_{010} = -R^1{}_{001} = \frac{4}{4t^2-1}$$

- Ricci tensor

$$R_{00} = \frac{4}{4t^2-1} \qquad R_{11} = -\frac{4t^2}{(4t^2-1)^2}$$

- Ricci Scalar

$$R = -\frac{8}{(4t^2-1)^2}.$$

All other components are zero. While the connection coefficients are singular at $t = 0$ and/or $t = 1/2$, the curvature tensors are only singular at $t = 1/2$. Furthermore these are not removable singularities.

[3]Admittedly, this is not a actually a tensor.

There is no need to calculate the Einstein tensor since it always vanishes in two dimensions (see Collas [1977]). Depending on your point of view the Einstein equations are either not-applicable or exactly satisfied (and exclude the possibility of matter). Either way it cannot be used to obtain any more information about the spacetime. Chapter 7 contains a generalization to higher dimensions, where the Einstein tensor does not vanish. The special case of four dimensions is used as a stage for a cosmological analysis of the model spacetime.

As far as general geodesics go, they raise an interesting question. How does the Riemannian geometry before the transition influence the behavior of particles after the transition? Or, turning the question on its head, is there any meaning that can be ascribed to geodesics in the Riemannian domain via their relationship to geodesics in the Lorentzian domain? These questions depend on the possibility of matching geodesic solutions across the transition surface. An investigation into whether or not geodesics on one domain can be matched to those of the other domain is relegated to Appendix A. In what follows, I will mostly consider the behavior of a Klein-Gordon field in the toy-model spacetime just described. While it is an interesting undertaking for its own sake, a Klein-Gordon field is a very powerful probe for a given spacetime. It can serve both as a proxy for geodesics and it is also a simple starting point for quantum mechanical investigations. Before delving into the analysis of a Klein-Gordon field, I discuss the connection between Klein-Gordon fields and geodesics, namely Hamilton's optico-mechanical analogy.

Chapter Three: Optico-mechanical Analogy and Geometric Optics

This chapter describes the relationship between a Klein-Gordon field and geodesics on spacetime. The link between them is the geometric optics approximation. This approximation is well known for linking the wave theory of light to ray optics, which was ultimately generalized by Hamilton into his Optico-mechanical analogy. The details of these relationships are the focus of this chapter, which serve as partial motivation for the work in this Thesis.

3.1 Classical Wave-Particle Duality of Light

Long before Maxwell's successful reformulation of light as an electromagnetic wave phenomena, Fermat had captured many of the properties of light (in particular reflection and refraction) by modeling light as rays described by curves/paths in space

$$\vec{r} : \mathbb{R} \to \mathbb{R}^3$$

subject to the **Fermat's Principle**. This principle states that the path a light ray takes between a point $\vec{q}$ and a point $\vec{p}$ in space (thought of as $\mathbb{R}^3$) is an extremum of the **optical length** integral

$$\int_{\vec{p}}^{\vec{q}} n(\vec{r}) dl$$

where $n : \mathbb{R}^3 \to \mathbb{R}$ is the index of refraction (which for the sake of simplicity is assumed to be isotropic) and $dl^2 = d\vec{r} \cdot d\vec{r}$ with $\cdot$ representing the Euclidean inner product. In other words light follows geodesics in space, with the Riemannian metric

$$g_{ij}(\vec{r}) = \delta_{ij} n(\vec{r})^2.$$

Plugging said metric into the non-affinely parameterized geodesic equation leads to

$$\frac{d^2 r^i}{d\lambda^2} + \frac{1}{n}\left(\delta^i_l \partial_k n + \delta^i_k \partial_l n - \delta^{ij}\delta_{kl}\partial_j n\right)\frac{dr^k}{d\lambda}\frac{dr^l}{d\lambda} = \frac{dr^i}{d\lambda}\frac{d}{d\lambda}\ln\left(n\sqrt{\delta_{kl}\frac{dr^k}{d\lambda}\frac{dr^l}{d\lambda}}\right).$$

Rewriting the above expression in vector notation leads to

$$\frac{d^2\vec{r}}{d\lambda^2} + \frac{2}{n}\left(\nabla n \cdot \frac{d\vec{r}}{d\lambda}\right)\frac{d\vec{r}}{d\lambda} - \frac{1}{n}\left(\frac{d\vec{r}}{d\lambda}\cdot\frac{d\vec{r}}{d\lambda}\right)\nabla n = \frac{d}{d\lambda}\ln\left(n\sqrt{\frac{d\vec{r}}{d\lambda}\cdot\frac{d\vec{r}}{d\lambda}}\right)\frac{d\vec{r}}{d\lambda},$$

where $\cdot$ is still being used to represent the Euclidean inner product. The equation above is not particularly recognizable in its current form. However, if instead of using an arbitrary parameterization λ, the (euclidean) arc-length parameter s defined by the condition

$$\frac{d\vec{r}}{ds}\cdot\frac{d\vec{r}}{ds} = 1$$

is chosen, the equation can be simplified to

$$\frac{d}{ds}\left(n\frac{d\vec{r}}{ds}\right) = \nabla n.$$

This equation is known as the **vector eikonal equation** and it governs the local behavior of light rays according to Fermat's principle.

Note that the parameter s is not an affine parameter with respect to the Riemannian metric $g_{ij} = \delta_{ij}n^2$. Instead, an affine parameter τ satisfies the condition

$$\frac{d\vec{r}}{d\tau}\cdot\frac{d\vec{r}}{d\tau} = \frac{1}{n^2}.$$

This is starting to look like the expression for the velocity of a ray of light, which is more evident if the affine parameter $t = \tau/c$ is used instead, where c is the speed of light in vacuum. Using the parameter t, which can be understood to be time, the equation governing the path $\vec{r}(t)$ becomes

$$\frac{d^2\vec{r}}{dt^2} + \frac{2}{n}\left(\nabla n \cdot \frac{d\vec{r}}{dt}\right)\frac{d\vec{r}}{dt} = \frac{c^2}{n^3}\nabla n.$$

At around the same time, there was an alternative characterization of light in terms of waves. It was qualitatively described by **Huygens principle**. Huygens proposed that light can be modeled as waves in space made up of propagating surfaces called **wavefronts**. In the modern understanding of waves, these wavefronts are simply surfaces of constant phase at a fixed instant in time. However, Huygens envisioned these waves as more like pulses than waves (see Complete Dictionary of Scientific Biography [2008]); he did not talk about periodicity or regularity.

According to the *Huygens principle*, given a wavefront at a fixed time t, the way to evolve that wavefront forward in time is to consider each point along the wavefront

to be a new source of light, emitting spherical waves, or **wavelets**. These wavelets travel at the speed determined by the index of refraction of the medium, analogous to Fermat's principle. At a later point in time, the new wavefront is variously described as the "sum of" or the "envelope of" these individual wavelets. One way to arrive at this envelope of wavelets, is to take each point in a wavefront (or a finite but dense subset thereof) and propagate it forwards along the direction normal to the wavefront, for a distance determined by the speed of the wave in the medium and the time of travel. For a variable index of refraction, this is only exactly accurate in the limit of infinitesimally small time steps[1].

Huygens principle, as stated, while providing great physical insight into the nature of light, is not very analytical. Obtaining the new wavefronts from old ones is more of an exercise in drafting. However Huygens was able to reproduce the laws of reflection and refraction with his principle. Over a hundred years later Fresnel independently discovered Huygens principle and expanded upon it by actually treating these wavelets as waves; including the concepts of phase, amplitude and interference[2]. With this addition he was able to also explain diffraction of light.

It is illustrative to derive a mathematical version of Huygens principle from Maxwell's equations, which came after both Huygens and Fresnel. In particular start with the source-free Maxwell's equations

$$\nabla \cdot \vec{D} = 0 \qquad \nabla \cdot \vec{B} = 0$$
$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \qquad \nabla \times \vec{H} = \frac{\partial \vec{D}}{\partial t}$$

with constituitive relations for isotropic, inhomogenoeous, linear media

$$\vec{D}(x,y,z,t) = \epsilon(x,y,z)\vec{E}(x,y,z,t) \qquad \vec{H}(x,y,z,t) = \frac{1}{\mu(x,y,z)}\vec{B}(x,y,z,t).$$

By taking another curl of the first equation, a time derivative of the second, plugging and manipulating vector expressions, the following wave equation for the components of the electric field can be obtained

$$\epsilon\mu\frac{\partial^2 \vec{E}}{\partial t^2} - \nabla^2\vec{E} = \left(\nabla\tilde{\epsilon}\cdot\nabla\right)\vec{E} + \left(\vec{E}\cdot\nabla\right)\nabla\tilde{\epsilon} + \nabla\left(\tilde{\epsilon}+\tilde{\mu}\right)\times\left(\nabla\times\vec{E}\right),$$

[1] It is interesting to note that Huygens principle predates calculus by a few years, and perhaps not so coincidentally Huygens taught Leibniz mathematics, when Leibniz was first getting into the subject.

[2] Fresnel had the distinct advantage over Huygens that the wave equation was introduced and solved by D'Alembert in the intervening period.

where $\tilde{\epsilon} = \ln(\epsilon)$ and $\tilde{\mu} = \ln(\mu)$. A similar equation can be obtained for the magnetic vector field (see Habib Mazharimousavi et al. [2013]). The next step is to plug in an ansatz of the form

$$\vec{E} = \vec{F}(x, y, z, t)e^{\frac{i}{\delta}\phi(x,y,z,t)}$$

where δ is a small parameter. This is often the first step when searching for an asymptotic series solution to an equation in a singular perturbation limit. This framework of asymptotic analysis will be discussed in chapter 8. In this case, the limit that is being considered is the limit in which the phase is both numerically much larger and varies much faster than the amplitude. This is often called the **geometric optics** limit. It is the limit in which the solution to the electromagnetic wave equation behaves similar to a monochromatic, plane-wave solution. Physically this limit corresponds to the situation where the length scale $\mathcal{L}$ and time scale $\mathcal{T}$ of the problem are long compared to the waves' period and wavelength. Plugging the ansatz above and keeping only the real part of the equation leads to

$$\epsilon\mu\frac{\partial^2\vec{F}}{\partial t^2} - \nabla^2\vec{F} - \frac{1}{\delta^2}\left(\epsilon\mu\left(\frac{\partial\phi}{\partial t}\right)^2 - \nabla\phi\cdot\nabla\phi\right) = \left(\nabla\tilde{\epsilon}\cdot\nabla\right)\vec{F} + \left(\vec{F}\cdot\nabla\right)\nabla\tilde{\epsilon} + \nabla\left(\tilde{\epsilon}+\tilde{\mu}\right)\times\left(\nabla\times\vec{F}\right),$$

which in the limit as $\delta \to 0$ becomes

$$\nabla\phi\cdot\nabla\phi = \epsilon\mu\left(\frac{\partial\phi}{\partial t}\right)^2.$$

This is one version of the **eikonal equation**. For the case of a plane monochromatic plane wave, where $\phi \propto (\vec{k}\cdot\vec{r} - \omega t)$, the eikonal equation reduces to the dispersion relation

$$k^2 = \frac{1}{v^2}\omega^2.$$

Furthermore, the eikonal equation can be interpreted as an infinitesimal expression of Huygens principle. Said principle is a statement about wavefronts, which are the equal phase surfaces at fixed time t_0. To find the wavefronts as a function of time, a solution to the eikonal equation is sought, in the form $\phi(x, y, z, t) = S(x, y, z) + T(t)$. Plugging this ansatz into the eikonal equation leads to a separation of variables

$$\frac{1}{\epsilon\mu}\nabla S\cdot\nabla S = \tilde{\omega}^2 = \dot{T}^2,$$

where $\tilde{\omega}^2$ is an arbitrary constant. The choice of notation, on the other hand, is not arbitrary, for $\tilde{\omega}$ is related to the angular frequency of the plane wave (in the geometric

optics limit). Indeed the approximate solution to the wave equation now takes the form

$$\vec{E} = \vec{F} e^{i\left(\frac{S(x,y,z)}{\delta} \pm \frac{\tilde{\omega}}{\delta} t\right)}$$

where $\vec{F}$ is approximately constant (relative to S/δ and $\tilde{\omega}/\delta$). So $\tilde{\omega}/\delta$ is effectively the angular frequency and the level surfaces of S are the wavefronts. Furthermore, letting $\epsilon\mu = n^2/c^2$, where n is the (potentially variable) index of refraction and c is the speed of light in vacuum, the equation governing S becomes

$$\nabla S \cdot \nabla S = \frac{n^2 \tilde{\omega}^2}{c^2}.$$

This equation is also known as the *scalar eikonal equation*. See Arley [1945] for an alternative and more detailed derivation of this form of the eikonal equation. One way to interpret the scalar eikonal equation is as follows. Pick a surface of constant S (i.e. a wave front). For any point $\vec{r}$ along this wavefront, find the new point

$$\vec{r}' = \vec{r} + \frac{\delta s}{n(\vec{r})^2} \nabla S|_{\vec{r}}$$

for a fixed real number δs. In other words, move a distance

$$\left| \frac{\delta s}{n(\vec{r})^2} \nabla S|_{\vec{r}} \right| = \frac{\tilde{\omega} \delta s}{n(\vec{r}) c}$$

in the direction normal to the wavefront at the point $\vec{r}$. The locus of points arrived to in this way will be another wavefront - up to higher order δs corrections. This can be seen from

$$\begin{aligned} S(\vec{r'}) &= S\left(\vec{r} + \frac{\delta s}{n(\vec{r})^2} \nabla S|_{\vec{r}}\right) \\ &\approx S(\vec{r}) + \frac{\delta s}{n(\vec{r})^2} \left(\nabla S \cdot \nabla S\right)|_{\vec{r}} \\ &= S(\vec{r}) + \frac{\tilde{\omega}^2 \delta s}{c^2}. \end{aligned}$$

The value of S at the new points $\vec{r}'$ only depends on the value of S at the points $\vec{r}$. So the points $\vec{r'}$ are also a wavefront. Notice that redefining δs as $c^2 \delta t/\tilde{\omega}$, the distance between $\vec{r}$ and $\vec{r}'$ as defined above becomes

$$v(\vec{r}) \delta t$$

which is in keeping with Huygens principle.

Now, this whole exercise of wave and particle approaches to light culminates with the realization that identifying the unit normal to the wavefronts of S - à la Huygens - with the unit tangent of the rays $\vec{r}$ - à la Fermat - namely

$$\frac{c}{\tilde{\omega}n}\nabla S = \frac{d\vec{r}}{ds},$$

ties both wave and particle perspectives together. On the one hand given a function S satisfying the scalar eikonal equation, letting $d\vec{r}/ds$ be defined according to the prescription above leads to

$$\begin{aligned}
\frac{d}{ds}\left(n\frac{d\vec{r}}{ds}\right) &= \frac{d\vec{r}}{ds}\cdot\nabla\left(\frac{c}{\tilde{\omega}}\nabla S\right) \\
&= \frac{c^2}{\tilde{\omega}^2 n}\nabla S\cdot\nabla\left(\nabla S\right) \\
&= \frac{c^2}{2\tilde{\omega}^2 n}\nabla\left(\nabla S\cdot\nabla S\right) \\
&= \frac{1}{2n}\nabla(n^2) = \nabla n
\end{aligned}$$

The result is precisely the vector eikonal equation for light rays derived from Fermat's principle.

On the other hand, Fermat's principle for light rays guarantees that given a solution of the vector eikonal equation in terms of a family of rays, there is a function S such that the tangents $d\vec{r}/ds$ and the gradient of this function S are related as above. Furthermore, this function will satisfy the scalar eikonal equation. That this is true is far from obvious, and will be demonstrated in a broader context in the following section. It turns out that the function S in question will be a reinterpretation of the action functional used to state Fermat's principle. Hence the potentially confusing choice of notation of using S to denote the phase function in Huygen's principle, on top of its typical use as an action functional. This duality was understood by Hamilton, and further developed into what is now known as Hamilton's optico-mechanical analogy, which is the subject of the next section.

3.2 Hamilton's Optico-Mechanical Analogy

While working in optics, Hamilton obtained a deep understanding of the dual descriptions of light: Fermat's rays and Huygens wave fronts. Furthermore, he realized that

this duality could be extended into the realm of mechanics, which was at the time solely described in terms of paths, or rays. This realization led to the development of Hamiltonian mechanics and the Hamilton-Jacobi theory. The point of departure was the encoding of the dynamical laws via the stationary action principle. This principle states that the the paths particles take are those that extremize the action functional

$$S[\gamma] = \int_{t_i}^{t_f} L(q^\mu, \dot{q}^\nu, t)|_\gamma dt.$$

In the modern understanding, S is a map from the (appropriate) space of paths in configuration space M to the real numbers, also known as a functional. Therefore the argument of S, here denoted γ, must be a path $\gamma : \mathbb{R} \to M$. The Lagrangian L is a real valued function on the (extended) tangent bundle of the configuration space $TM \times \mathbb{R}$, the coordinates of which are denoted $(q^\mu, \dot{q}^\mu, t)$. To evaluate the integral, the Lagrangian function is evaluated on the (lift of the) path γ.

The actual path taken by a particle is that which extremizes the action. Without too many details, the approach to deriving the condition satisfied by classical path γ_c is to consider a family of paths around the classical path

$$\gamma_c + \delta\gamma$$

such that the endpoints are fixed. That means that all the paths start from a fixed point q_i at time t_i and end at another fixed point q_f at time t_f. In particular this implies that $\delta\gamma(t_i) = \delta\gamma(t_f) = 0$. The difference in action between the path γ_c and the paths $\gamma_c + \delta\gamma$, up to first order in the variation $\delta\gamma$ is

$$S[\gamma_c + \delta\gamma] - S[\gamma_c] = \int_{t_i}^{t_f} dt \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^\mu} \delta\gamma^\mu \right) + \left(\frac{\partial L}{\partial q^\mu} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^\mu} \right) \delta\gamma^\mu \right) + O(\delta\gamma^2),$$

where the partial derivatives of the Lagrangian are all evaluated on the path γ_c, and the $\delta\gamma^\mu$ are the component functions of the variation $\delta\gamma$. The total time derivative term can be immediately integrated

$$\int_{t_i}^{t_f} dt \; \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^\mu} \delta\gamma^\mu \right) = \left(\frac{\partial L}{\partial \dot{q}^\mu} \delta\gamma^\mu \right)\Big|_{t_i}^{t_f},$$

which is equal to zero since $\delta\gamma$ is zero at t_i and t_f. The remaining piece in the term linear in $\delta\gamma$ vanishes when

$$\left.\frac{\partial L}{\partial q^\mu}\right|_{\gamma_c} - \frac{d}{dt}\left(\left.\frac{\partial L}{\partial \dot{q}^\mu}\right|_{\gamma_c} \right) = 0.$$

These are the Euler-Lagrange equations; they are the condition satisfied by paths that extremize the action.

Now, the *action functional* can be used to define an *action function*. This is a function from the configuration space cross the time axis $M \times \mathbb{R}$, to the real numbers. This function is denoted $S(q,t)$, as opposed to the action functional $S[\gamma]$. The idea is to fix an initial point (q_i, t_i) for the action integral and let the endpoint (q,t) vary with the understanding the the integration will be performed along the extremal curve connecting the initial point to the point (q,t). I will use the following notation for this function:

$$S(q,t) = \int_{t_i}^{t} L(q^\mu, \dot{q}^\nu, t)|_{\gamma_{(q,t)}}\, dt$$

where the subscript (q,t) on the curve $\gamma_{(q,t)}$ is denoting that the extremal curve being integrated over is the one starting at (q_i, t_i) and ending at (q,t). It is not clear that this is a well defined function. Indeed further care must be taken to properly define this function. For a discussion of the subtleties and how to resolve them see the excellent treatment by Arnold [1989] in ch. 8.

It is this action function that takes the place of Huygens phase function in the optical realm. This can be seen by computing both the space and time derivatives of $S(q,t)$. First, to get the space derivative, consider a small displacement $q + \delta q$. The extremum curve $\gamma_{(q+\delta q,t)}$ that goes from (q_i, t_i) to $(q + \delta q, t)$ will be different than the extremum curve $\gamma_{(q,t)}$ that goes from (q_i, t_i) to (q,t). For small enough displacement however, the curve $\gamma_{(q+\delta q,t)}$ can be written as[3]

$$\gamma_{(q+\delta q,t)} = \gamma_{(q,t)} + \delta\gamma$$

where $\delta\gamma$ is small, $\delta\gamma(t_i) = 0$ and $\delta\gamma(t) = \delta q$. In what follows, let $\gamma_{(q,t)}$ be denoted simply as γ to keep the equations from overflowing. The derivative can therefore be computed by taking the difference between $S(q + \delta q, t)$ and $S(q,t)$, and keeping only

[3]The notation $\gamma + \delta\gamma$ sacrifices precision for the sake of simplicity. Strictly speaking γ cannot be added to $\delta\gamma$ since the manifold M does not have an additive group structure. However, in coordinates, they can be added since $\mathbb{R}^n$ does have such a structure. So the addition should be understood to happen after the paths have been expressed in coordinates.

the term linear in δq:

$$\begin{aligned}
\Delta S &= \int_{t_i}^{t} L|_{\gamma+\delta\gamma}\, dt - \int_{t_i}^{t} L|_{\gamma}\, dt \\
&= \int_{t_i}^{t} \left(L|_{\gamma} + \frac{\partial L}{\partial q^\mu}\bigg|_{\gamma} \delta\gamma^\mu + \frac{\partial L}{\partial \dot{q}^\mu}\bigg|_{\gamma} \frac{d}{dt}\delta\gamma^\mu \right) dt - \int_{t_i}^{t} L|_{\gamma}\, dt + O(\delta\gamma^2) \\
&= \int_{t_i}^{t} \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^\mu}\bigg|_{\gamma} \delta\gamma^\mu \right) + \left(\frac{\partial L}{\partial q^\mu}\bigg|_{\gamma} + \frac{d}{dt}\, \frac{\partial L}{\partial \dot{q}^\mu}\bigg|_{\gamma} \right) \delta\gamma^\mu \right) dt + O(\delta\gamma^2).
\end{aligned}$$

The second term in the integral is precisely the Euler Lagrange equations being evaluated on the path γ. Since γ is an extremal curve, this piece vanishes. On the other hand, the first term, which is a total derivative, does not vanish. It can be integrated to obtain

$$\int_{t_i}^{t} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^\mu}\bigg|_{\gamma} \delta\gamma^\mu \right) dt = \left(\frac{\partial L}{\partial \dot{q}^\mu}\bigg|_{\gamma} \delta\gamma^\mu \right)\bigg|_{t_i}^{t}.$$

While $\delta\gamma(t_i) = 0$, the variation at the other endpoint is not equal to zero, instead $\delta\gamma(t) = \delta q$. So the derivative of S with respect to q becomes

$$\frac{\partial S}{\partial q^\mu} = \frac{\partial L}{\partial \dot{q}^\mu} \tag{3.1}$$

evaluated on path γ at time t.

To evaluate the derivative of the action function $S(q,t)$ with respect to time requires a similar procedure; that is take the difference between $S(q, t+\delta t)$ and $S(q,t)$ and look at the term linear in δt. Even though q is not changing, the extremum curve $\gamma_{(q,t+\delta t)}$ is still a different one than $\gamma_{(q,t)}$, since it must "arrive" at q at a different time. So a variation in curves must still be considered. Once again, let

$$\gamma_{(q,t+\delta t)} = \gamma + \delta\gamma$$

where $\delta\gamma(t_i) = 0$, but this time the value of $\delta\gamma$ at the other endpoint $t + \delta t$ is a little different. In order to be clear, consider the component functions of $\gamma_{(q,t+\delta t)}$, γ and $\delta\gamma$, in the coordinates $(q^\mu, \dot{q}^\mu, t)$. These are denoted $\gamma^\mu_{(q,t+\delta t)}$, γ^μ and $\delta\gamma^\mu$, respectively.

The equation relating these functions is

$$\begin{aligned}\delta\gamma^\mu(t+\delta t) &= \gamma^\mu_{(q,t+\delta t)}(t+\delta t) - \gamma^\mu(t+\delta t)\\ &\approx q^\mu - \left(\gamma^\mu(t) + \left.\frac{d\gamma^\mu}{dt}\right|_t \delta t\right)\\ &= -\left.\frac{d\gamma^\mu}{dt}\right|_t \delta t\end{aligned}$$

where in getting from the second to the third line, the fact that $\gamma^\mu(t) = q^\mu$ was used. The term $d\gamma^\mu/dt$ represents the components of the velocity vector γ' at time t, in the (q^μ) coordinates. It can also, however, be interpreted as the value of the $\dot{q}^\mu$ coordinate of the point on $TM \times \mathbb{R}$ which the lift of the curve γ evaluates to, at time t. Therefore, the difference between $S(q, t+\delta t)$ and $S(q,t)$ can be expressed as follows:

$$\begin{aligned}\Delta S &= \int_{t_i}^{t+\delta t} L|_{\gamma+\delta\gamma}\, dt - \int_{t_i}^{t} L|_\gamma\, dt\\ &= \int_{t_i}^{t+\delta t}\left(L|_\gamma + \left.\frac{\partial L}{\partial q^\mu}\right|_\gamma \delta\gamma^\mu + \left.\frac{\partial L}{\partial \dot{q}^\mu}\right|_\gamma \frac{d}{dt}\delta\gamma^\mu\right) dt - \int_{t_i}^{t} L|_\gamma\, dt + O(\delta\gamma^2)\\ &= \int_{t_i}^{t+\delta t}\frac{d}{dt}\left(\left.\frac{\partial L}{\partial \dot{q}^\mu}\right|_\gamma \delta\gamma^\mu\right) dt + \int_{t_i}^{t+\delta t} L|_\gamma dt - \int_{t_i}^{t} L|_\gamma dt + O(\delta\gamma^2)\\ &= \left.\left(\left.\frac{\partial L}{\partial \dot{q}^\mu}\right|_\gamma \delta\gamma^\mu\right)\right|_{t+\delta t} + \int_{t}^{t+\delta t} L|_\gamma dt + O(\delta\gamma^2).\end{aligned}$$

The first of these terms becomes

$$\left.\left(\left.\frac{\partial L}{\partial \dot{q}^\mu}\right|_\gamma \delta\gamma^\mu\right)\right|_{t+\delta t} = -\left.\left(\left.\frac{\partial L}{\partial \dot{q}^\mu}\right|_\gamma\right)\right|_{t+\delta t}\left.\frac{d\gamma^\mu}{dt}\right|_t \delta t \approx -\left.\left(\left(\frac{\partial L}{\partial \dot{q}^\mu}\dot{q}^\mu\right)\right|_\gamma\right)\right|_t \delta t,$$

while the second one becomes

$$\int_t^{t+\delta t} L|_\gamma dt = (L|_\gamma)|_t\, \delta t.$$

All together then, from the difference between $S(q, t+\delta t)$ and $S(q,t)$, up to terms linear in order of δt, the derivative $\frac{\partial S}{\partial t}$ is

$$\frac{\partial S}{\partial t} = -\left.\left(\frac{\partial L}{\partial \dot{q}^\mu}\dot{q}^\mu - L\right)\right|_t, \tag{3.2}$$

where L, its partial derivative, and $\dot{q}^\mu$ are evaluated on the extremum path γ.

In the context of Hamiltonian Mechanics, a theory developed around these previous considerations, the derivatives of the action function $S(q,t)$ have special significance. The derivative with respect to generalized coordinates is equal to $\partial L/\partial \dot{q}^\mu$. As a function of on $TM \times \mathbb{R}$, this is relabeled as

$$p^\mu(q,\dot{q},t) := \frac{\partial L}{\partial \dot{q}^\mu}$$

and is called the *conjugate momentum function*. On the other hand, the derivative with respect to time equals $\sum \dot{q}^\mu p^\mu - L$, which is the Legendre transform of L, when replacing $\dot{q}^\mu$ with the conjugate momenta $p^\mu = \frac{\partial L}{\partial \dot{q}^\mu}$ defined above. In particular

$$H(q^\nu, p^\nu, t) := \sum_\mu p^\mu \dot{q}^\mu(q^\nu, p^\nu, t) - L(q^\nu, \dot{q}^\nu(q^\nu, p^\nu, t), t),$$

and is called the *Hamiltonian*. This procedure can only be carried out if the relationship $p^\mu = p^\mu(q^\nu, \dot{q}^\nu, t)$ can be inverted into a relationship of the form

$$\dot{q}^\mu = \dot{q}^\mu(q^\nu, p^\nu, t),$$

in which case the Lagrangian is labeled as *non-singular* or *non-degenerate*. If such condition holds, then putting these results together, the following equation for the action function $S(q,t)$ is obtained:

$$\frac{\partial S}{\partial t} + H\left(q^\nu, \frac{\partial S}{\partial q^\nu}, t\right) = 0.$$

This is known as the Hamilton-Jacobi equation. An incredibly rich theory of canonical transformations has been developed around it (see Arnold [1989]).

For the purpose at hand, the result of this investigation is the following. Given a mechanical theory of particles, specified by a extremal action principle with a Lagrangian L, the canonical momentum $p(q,\dot{q},t)$ and Hamiltonian $H(q,p,t)$ can be computed, provided the definition of conjugate momenta can be inverted into a function $\dot{q}(q,p,t)$. Furthermore, if a function $S(q,t)$ is found, which satisfies the associated Hamilton-Jacobi equation

$$\frac{\partial S}{\partial t} + H\left(q^\nu, \frac{\partial S}{\partial q^\nu}, t\right) = 0,$$

then the integral curves to the vector field

$$\dot{q}^\mu(q^\nu, \nabla S, t)$$

will satisfy the Euler Largange equations of motion.

This can be seen as follows

1. ∇S is equal to p, the canonical momentum, evaluated on the path that extremizes the action functional - that is the path that solves the Euler-Lagrange equations.

2. Equation 3.1 can be interpreted as

$$p(q,t) = \frac{\partial L}{\partial \dot{q}}(q, \dot{q}, t).$$

   This is equivalent (given non-degeneracy) to the specification of $\dot{q}$ as a function of (q,t) - as determined by the Euler-Lagrange equations. This is simply a vector field, whose integral curves give rise to paths, and they will satisfy Euler-Largange.

This is just a re-intepretation of the condition

$$\left.\frac{\partial L}{\partial \dot{q}^\mu}\right|_\gamma = \nabla S$$

using the expression for $\dot{q}^\mu$ in terms of (q,p,t). Thus the dual description of light, in terms of paths and real valued functions has been imported into the realm of mechanics.

For a quick example consider a free, non relativistic particle in two dimensions, using cartesian coordinates (x,y). The Lagrangian, conjugate momentum and, Hamiltonian are respectively

$$L = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right) \qquad (p_x, p_y) = m(\dot{x}, \dot{y}) \qquad H = \frac{1}{2m}\left(p_x^2 + p_y^2\right).$$

Therefore the Hamilton Jacobi equation is

$$\frac{\partial S}{\partial t} + \frac{1}{2m}\left(\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2\right) = 0.$$

On the other hand, the Euler Largange equations from the extremum action principle are

$$m\ddot{x} = 0 \qquad m\ddot{y} = 0$$

whose solutions, for a path starting from (x_i, y_i) at t_i and ending at (x_f, y_f) at t_f, take the form

$$x = \frac{(x_f - x_i)}{(t_f - t_i)}(t - t_i) + x_i \qquad y = \frac{(y_f - y_i)}{(t_f - t_i)}(t - t_i) + y_i. \tag{3.3}$$

Plugging these paths back into the action results in

$$\begin{aligned} S &= \int_{t_i}^{t_f} \frac{1}{2} m \left(\left(\frac{(x_f - x_i)}{(t_f - t_i)} \right)^2 + \left(\frac{(y_f - y_i)}{(t_f - t_i)} \right)^2 \right) dt \\ &= \frac{1}{2} m \left(\frac{(x_f - x_i)^2}{(t_f - t_i)} + \frac{(y_f - y_i)^2}{(t_f - t_i)} \right), \end{aligned}$$

which after relabeling the endpoint (x_f, y_f, t_f) as simply (x, y, t), results in the action function

$$S(x, y, t) = \frac{1}{2} m \left(\frac{(x - x_i)^2}{(t - t_i)} + \frac{(y - y_i)^2}{(t - t_i)} \right).$$

A quick calculation shows that this function does indeed satisfy the Hamilton Jacobi equation.

$$\begin{aligned} \frac{1}{2m} \left(\left(\frac{\partial S}{\partial x} \right)^2 + \left(\frac{\partial S}{\partial y} \right)^2 \right) &= \frac{1}{2} m \left(\left(\frac{(x - x_i)}{(t - t_i)} \right)^2 + \left(\frac{(y - y_i)}{(t - t_i)} \right)^2 \right) \\ &= -\frac{\partial S}{\partial t}. \end{aligned}$$

Furthermore, the paths taken by particles can be recovered through the fact that they will be integral curves of the vector field defined as

$$\dot{q}^\mu(q^\nu, \nabla S, t),$$

Solving for the velocities in terms of the momentum, to obtain the function $\dot{q}^\mu = \dot{q}^\mu(q^\nu, p^\mu, t)$ yields

$$(\dot{x}, \dot{y}) = \frac{1}{m}(p_x, p_y). \tag{3.4}$$

Then plugging in the values

$$\begin{aligned} p_x &= \partial_x S & (3.5) \\ &= m\frac{(x - x_i)}{(t - t_i)} & (3.6) \end{aligned}$$

and

$$p_y = \partial_y S \tag{3.7}$$
$$= m\frac{(y-y_i)}{(t-t_i)} \tag{3.8}$$

leads to the vector field

$$\frac{(x-x_i)}{(t-t_i)}\hat{e}_x + \frac{(x-y_i)}{(t-t_i)}\hat{e}_y,$$

where $\hat{e}_x$ and $\hat{e}_y$ are unit vectors in the x and y directions respectively. The integral curves to this vector field are indeed the curves in equation 3.3.

To finalize this section, it is instructive to go back and check the final claim made in section (3.1). That is, the description of light in term of paths, via Fermat's principle, is ultimately equivalent to that in terms of wavefronts, via Huygens' principle. It was already shown that given a solution to the scalar eikonal equation, a congruence of paths could be defined such that they satisfied the vector eikonal equation. All that remains to be shown is the converse. In particular that given paths satisfying Fermat's principle, there exists a function, satisfying the scalar eikonal equation, such that the normal to the level curves (at a fixed time t) give rise a vector field tangent to the ray paths. The entire re-interpretation of the action functional into an action function was done with this goal in mind. So presumably the next step is to write down the Hamilton-Jacobi equation associated with the functional in Fermat's principle. There is a complication however, the conjugate momentum defined by the Lagrangian in Fermat's principle cannot be inverted into an expression for the generalized velocity coordinate as a function of the conjugate momentum. The solution is to consider an alternative, but in a sense equivalent action. In particular consider the action

$$S = \int d\lambda \left(\frac{1}{2}n^2(\vec{r})\frac{d\vec{r}}{d\lambda}\cdot\frac{d\vec{r}}{d\lambda}\right).$$

The Euler-Lagrange equation associated with this action is

$$\frac{d^2\vec{r}}{d\lambda^2} + \frac{2}{n}\left(\nabla n\cdot\frac{d\vec{r}}{d\lambda}\right)\frac{d\vec{r}}{d\lambda} - \left(\frac{d\vec{r}}{d\lambda}\cdot\frac{d\vec{r}}{d\lambda}\right)\frac{\nabla n}{n} = 0$$

which are the same as the equations derived from Fermat's principle, with affine parameter. The conjugate momentum and Hamiltonian are

$$\vec{p} = n^2\frac{d\vec{r}}{d\lambda} \qquad H = \frac{1}{2n^2}\vec{p}\cdot\vec{p}.$$

This leads to the Hamilton-Jacobi equation

$$\frac{\partial S}{\partial t} + \frac{1}{n^2}\nabla S \cdot \nabla S = 0.$$

To obtain the equation for the wavefronts it is necessary to isolate the dependence of S on t. Since the Hamiltonian is time independent, a separation of variable ansatz for S will lead to a simplification of the Hamilton-Jacobi equation, namely, let

$$S = \tilde{S}(\vec{r}) + T(t).$$

Plugging this into the Hamilton-Jacobi equation reduced to the following two equations

$$\dot{T} = -\alpha,$$

where α is a constant, and

$$\nabla \tilde{S} \cdot \nabla \tilde{S} = n^2 \alpha.$$

This second equation is precisely the scalar eikonal equation for the wavefronts (up to the interpretation of the constant α), as was to be shown. Furthermore, the wavefronts and rays are connected via the equation

$$\nabla \tilde{S} = n^2(\vec{r})\frac{d\vec{r}}{d\lambda}$$

as before (up to a constant).

3.3 Geodesics and the Klein-Gordon Equation

The Hamilton optico-mechanical analogy can be taken further. In particular, just like Huygen's wavefront equation arises as the geometric optics limit of Maxwell's electromagnetic wave equation, the Hamilton-Jacobi equation can be thought of a particular limit of some appropriate wave equation. In fact this is precisely how Schrödinger arrived at his celebrated equation (see Schrödinger [1926]). Shcrödinger's point of departure was de Broglie's idea that perhaps matter particles are also waves; a duality that had become evident in the context of light. He started from a classical particle in a conservative potential and used the optico-mechanical analogy to obtain an equation for the wavefronts of de Broglie's dual description. He then sought a wave equation that in the limit of small wavelength and period relative to classical scales, would yield the equation for the wave-fronts.

Similarly, given a semi-Riemannian manifold, a wave equation can be found such that, in the geometric optics limit, it results in wavefronts perpendicular to geodesics. Not surprisingly, the Klein-Gordon equation

$$g^{\mu\nu}\nabla_\mu\nabla_\nu\psi = \mu^2\psi$$

satisfies this requirement. This application of the optico-mechanical analogy follows from the characterization of geodesics in terms of an action principle. However, the typical action used in the definition of geodesics, namely the length integral

$$S[\gamma] = \int_{\lambda_0}^{\lambda_1} d\lambda \sqrt{(sgn) g_{\mu\nu} \frac{d\gamma^\mu}{d\lambda}\frac{d\gamma^\nu}{d\lambda}\Big|_\lambda},$$

does not work. As discussed in Appendix B, this action leads to non-invertible conjugate momenta and a vanishing Hamiltonian. On the other hand, the alternative action

$$S_2[\gamma^\mu] = \int_{\lambda_0}^{\lambda_1} d\lambda \ \left(\frac{1}{2} g_{\mu\nu}\frac{d\gamma^\mu}{d\lambda}\frac{d\gamma^\nu}{d\lambda}\right),$$

was shown to be equivalent, in the sense that the extremization condition also leads to the geodesic equation, with the caveat that the geodesic must be affinely parameterized. For this choice of action, the conjugate momenta and Hamiltonian are

$$p_\alpha = g_{\alpha\nu}\frac{d\gamma^\nu}{d\lambda} \qquad H(\gamma^\mu, p_\nu) = \frac{1}{2} g^{\alpha\beta} p_\alpha p_\beta$$

respectively. This leads to the Hamilton-Jacobi equation

$$\frac{\partial S}{\partial \lambda} + \frac{1}{2} g^{\alpha\beta}\frac{\partial S}{\partial x^\alpha}\frac{\partial S}{\partial x^\beta} = 0$$

for the associated phase function S. Using the ansatz $S = f(\lambda) + \tilde{S}(x^\mu)$ to look at the constant λ surfaces, the equation for S can be split into

$$S = m\lambda + \tilde{S} \qquad \frac{1}{2} g^{\alpha\beta}\frac{\partial \tilde{S}}{\partial x^\alpha}\frac{\partial \tilde{S}}{\partial x^\beta} = -m,$$

where m is a constant of integration.

On the other hand, starting from the Klein-Gordon equation and using the ansatz solution

$$\psi = A(x^\mu) e^{\frac{i}{\epsilon} B(x^\mu)}$$

where A and B are real valued functions, leads to the following equations for A and B, which are simply the real and imaginary parts (respectively) resulting from the substitution:

$$\frac{1}{\epsilon^2} g^{\mu\nu}\partial_\mu B\partial_\nu B - g^{\mu\nu}\frac{\partial_\mu A\partial_\nu A}{A} + g^{\mu\nu}\Gamma^\rho{}_{\mu\nu}\frac{\partial_\rho A}{A} + \mu^2 = 0$$
$$g^{\mu\nu}\partial_\mu\partial_\nu B - g^{\mu\nu}\Gamma^\rho{}_{\mu\nu}\partial_\rho B + 2g^{\mu\nu}\frac{\partial_\mu B\partial_\nu A}{A} = 0$$

In the above equations and the ones to follow, the notation $\partial_\mu A$ is used to represent the partial derivative of A with respect to x^μ. In the limit that ϵ goes to zero, the real part of the equation becomes

$$g^{\mu\nu}\partial_\mu B\partial_\nu B = 0.$$

This coincides with the Hamilton-Jacobi equation associated to the geodesics when the constant of integration m is set to zero, under the identification $\tilde{S} = B$. This in turn corresponds to a particle along a geodesic whose velocity vector satisfies

$$g_{\mu\nu}\frac{d\gamma^\mu}{d\lambda}\frac{d\gamma^\nu}{d\lambda} = 0,$$

i.e. a null geodesic.

So, the Klein-Gordon equation, in the limit of geometric optics (large phase and rapid phase change) reproduces the Hamilton-Jacobi equation for wavefronts dual to geodesic motion of particles, albeit apparently only for massless ones. This last fact can be remedied by being a little more careful when taking the geometric optics limit. In particular, if the assumption that as ϵ goes to zero, the constant μ^2 gets very large. Alternatively, the limit in which the phase, and phase change is much larger than the amplitude and amplitude change, but it is of about the same size as μ^2. In this case the real part of the Klein-Gordon equation with the geometric optics ansatz becomes

$$g^{\mu\nu}\partial_\mu B\partial_\nu B = -\epsilon\mu^2,$$

which thus allows for massive, time like particles in the analogy.

Chapter Four: Klein-Gordon - Approach 1

This chapter contains a first approach at understanding the behavior of a Klein-Gordon field across the transition surface. In the simple $n = 2$ and $\lambda = 0$ case, the Klein-Gordon equation can be solved exactly. Once the analytic solutions have been obtained for both the Riemannian and Lorentzian domain, a simple matching at the transition is attempted.

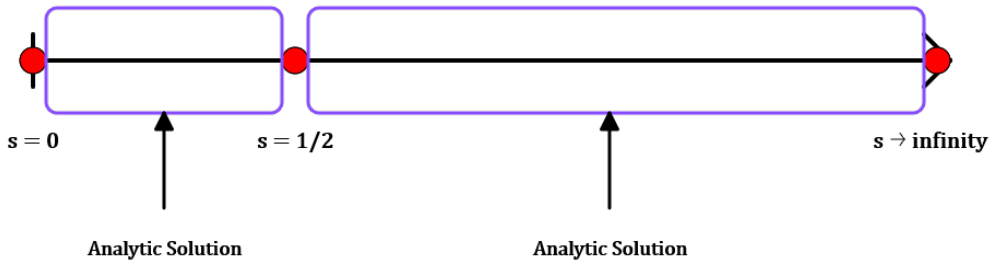


Figure 4.1: This approach uses analytic solutions in both domains.

4.1 Klein-Gordon: Initial Exploration

Before diving into the Klein-Gordon equation in the model transition-spacetime it is important to go back to the model definition and include the relevant unitful constants. Such constants set the scales for the problem, and allow the evaluation of the geometric optics limits in a meaningful way. The first step is to replace the Minkowski metric in the ambient three dimensional spacetime in which the paraboloid was embedded with the metric

$$-d(cT) \otimes d(cT) + d\rho \otimes d\rho + \rho^2 d\theta \otimes d\theta$$

where c is the speed of light constant, T has units of time, ρ has units of distance and θ is unitless. Furthermore, the condition that defines the embedded paraboloid

is replaced by

$$(cT) = \frac{\rho^2}{f}$$

where f is a constant with units of length; it is related to the focal length of the paraboloid, setting the scale of the curvature. The induced metric on the paraboloid becomes

$$-\left(\frac{4\rho^2}{f^2} - 1\right) d\rho \otimes d\rho + \rho^2 d\theta \otimes d\theta,$$

or instead, changing ρ to ct, it becomes

$$-\left(\frac{4(ct)^2}{f^2} - 1\right) d(ct) \otimes d(ct) + (ct)^2 d\theta \otimes d\theta.$$

It is important to keep in mind that in this (t, θ) coordinate system the speed of light is not c, rather it is

$$\frac{d\theta}{dt} = \pm\frac{1}{t}\sqrt{\left(\frac{2ct}{f}\right)^2 - 1}.$$

This is simply a restatement of the fact that these are not the coordinates of a freely falling observer. However, the fundamental constant c still has the objective meaning of being the speed at which light travels in locally inertial frames, that is, the local frames of freely falling observers.

Now, the Klein-Gordon equation for a scalar field Ψ follows from the action

$$S_{KG} = \alpha \int \left(-\frac{1}{2} g^{\mu\nu} \partial_\mu \Psi \partial_\nu \Psi - \frac{1}{2}\frac{\mu^2 c^2}{\hbar^2}\Psi^2\right) \sqrt{|g|}\, d(ct) \wedge d\theta.$$

The constant μ has units of mass, which therefore make $\mu^2 c^2/\hbar^2$ have units of inverse length squared (L^{-2}), which matches the kinetic piece[1]. The overall constant α is there to make the ensure that the action actually has units of action, that is energy times time. Whatever the units of the scalar field, the units of the product $\alpha\Psi^2$ must that of action:

$$[\alpha\Psi^2] = ET$$

While this constant α is irrelevant for the considerations of this section, it will become important in Ch 11, when the quantum behavior of the Klein-Gordon field is analyzed.

[1]The units of $g_{\mu\nu}$ depend on the value of the indices, i.e. g_{00} is unitless while g_{11} has units of length. However, when contracted with the appropriate coordinate derivative, the units do match.

The Klein-Gordon equation follows from this action via the typical extremization procedure (holding the metric components fixed), which results in

$$\frac{1}{\sqrt{-g}}\partial_\mu\left(\sqrt{-g}g^{\mu\nu}\partial_\nu\Psi\right)=\frac{\mu^2c^2}{\hbar^2}\Psi,$$

or equivalently

$$g^{\mu\nu}\nabla_\mu\nabla_\nu\Psi=\frac{\mu^2c^2}{\hbar^2}\Psi.$$

For the particular spacetime under consideration, the Klein-Gordon equation becomes

$$\frac{1}{\left(\frac{4(ct)^2}{f^2}-1\right)}\left(\frac{1}{c^2}\Psi_{tt}-\frac{1}{ct\left(\frac{4(ct)^2}{f^2}-1\right)}\frac{1}{c}\Psi_t\right)-\frac{1}{(ct)^2}\Psi_{\theta\theta}=-\frac{\mu^2c^2}{\hbar^2}\Psi.$$

To simplify the analysis, and make clear the limits that will be considered, the first step is to nondimensionalize the equation. In particular let

$$s=\frac{ct}{f}\qquad\Psi(t,\theta)=\psi(s(t),\theta).$$

The new variable s is unitless, and it counts how much time evolves in units of f/c. The Klein-Gordon equation becomes

$$\frac{1}{(4s^2-1)}\left(\psi_{ss}-\frac{1}{s\,(4s^2-1)}\psi_s\right)-\frac{1}{s^2}\psi_{\theta\theta}=-\lambda^2\psi \tag{4.1}$$

where λ is the unitless parameter

$$\lambda=\frac{f\mu c}{\hbar}.$$

In reference to section 3.3, a quick consistency check can be performed. Plugging an ansatz of the form $\psi=A(s,\theta)e^{\frac{i}{\epsilon}B(s,\theta)}$ into equation (4.1) above and keeping only the real part of the equation and the highest order terms of that, leads to

$$\frac{1}{\epsilon^2}\left(-\frac{B_s^2}{4s^2-1}+\frac{B_\theta^2}{s^2}\right)=-\lambda^2.$$

This is precisely

$$g^{\mu\nu}\partial_\mu B\partial_\nu B=-\epsilon^2\frac{\mu^2c^2}{\hbar^2}$$

as expected.

So, now the goal is to solve the Klein-Gordon equation and analyze the behavior at the transition boundary. Not only is this interesting in its own right, but it might be able to shine light on the behavior of geodesics in the appropriate limit. As a first step, a solution of the form $\psi(s,\theta) = S(s)\Theta(\theta)$ is sought. Because the equation is linear, any linear combination of solutions of this form is also a solution to the equation. In fact, solutions of this form, make a basis of solutions for the Klein-Gordon equation, so there is no loss of generality. Plugging this separation of variables ansatz into the Klein-Gordon, equation (4.1), leads to

$$\frac{s^2}{4s^2-1}\left(\frac{S''}{S} - \frac{1}{s(4s^2-1)}\frac{S'}{S}\right) + \lambda^2 s^2 = \frac{\Theta''}{\Theta}.$$

This implies that both the left hand side and the right hand side must be equal to a constant, temporarily denoted as β. The Klein-Gordon equation is thus separated into two ordinary differential equations

$$\begin{aligned} \Theta'' - \beta\Theta &= 0 \\ S'' - \frac{1}{s(4s^2-1)}S' + \frac{(-\beta+\lambda^2 s^2)(4s^2-1)}{s^2}S &= 0. \end{aligned}$$

The topology of the paraboloid, in particular the fact that the θ coordinate is parameterizing a circle, means that the constant β must have the form $-k^2$, where k is an integer. Otherwise the solutions to the Θ equation are not be periodic. Therefore the equations can be re-written as

$$\begin{aligned} \Theta'' + k^2\Theta &= 0 \\ S'' - \frac{1}{s(4s^2-1)}S' + \frac{(k^2+\lambda^2 s^2)(4s^2-1)}{s^2}S &= 0. \end{aligned}$$

As previously alluded to, the solutions to the Θ equation are of the form

$$\Theta = e^{ik\theta}, e^{-ik\theta}$$

or linear combinations thereof. As to the equation for S, lamentably it cannot be solved analytically for generic λ. However for the case of $\lambda = 0$ (which is $\mu = 0$), the equation can indeed be solved. To achieve this, perform the change of variables defined by the condition

$$du = \frac{\sqrt{4s^2-1}}{s}ds,$$

which can be integrated to obtain

$$u = \sqrt{4s^2 - 1} - \tan^{-1}\left(\sqrt{4s^2 - 1}\right).$$

For the moment this is purely a pragmatic coordinate transformation that simplifies obtaining the solution and not a statement about coordinates on the model spacetime.

Letting $U(u(s)) = S(s)$ the equation for S becomes

$$\frac{d^2U}{du^2} + (k^2 + \lambda^2 s(u)^2)U = 0.$$

This doesn't seem like much of an improvement, given that the equation defining u in terms of s cannot be inverted to obtain s in terms of u. However, in the case where $\lambda = 0$, the equation reduces to

$$U'' + k^2 U = 0,$$

whose solutions are of the form

$$U = e^{iku}, e^{-iku} \tag{4.2}$$

(or linear combinations thereof). Plugging back in the expression for u in terms of s results in the following linearly independent solutions for the equation governing S in the case of $\lambda = 0$:

$$S_1 = e^{ik\left(\sqrt{4s^2-1} - \tan^{-1}\left(\sqrt{4s^2-1}\right)\right)} \qquad S_2 = e^{-ik\left(\sqrt{4s^2-1} - \tan^{-1}\left(\sqrt{4s^2-1}\right)\right)}.$$

For $s > 1/2$, these are oscillatory functions with a time dependent frequency. Before moving on to a careful investigation of the solutions to the Klein-Gordon equation, consider the following solution

$$\psi(s, \theta) = e^{ik\left(\sqrt{4s^2-1} - \tan^{-1}\left(\sqrt{4s^2-1}\right) - \theta\right)}$$

which is simply constructed by multiplying the S_1 solution of the equation for S, and the negative exponential solution for the equation for Θ. This is an exact solution that closely resembles a monochromatic plane wave - in 1+1 dimensions - for $s > 1/2$. The constant k determines the number of wave cycles in space, which has the topology of a circle. For this solution, there is no need to take any geometric optics limit, since it is already effectively a plane wave, which is the prototypical behavior sought in the approximation.

To construct the dual description in terms of rays, the first step is to obtain the exterior derivative of the phase function labeled ϕ:

$$d\phi = \frac{\sqrt{4s^2 - 1}}{s} ds - d\theta.$$

According to the optico-mechanical analogy, this corresponds to the conjugate momentum to the particle trajectories. For the Hamiltonian used to define the geodesics (see Appendix B), the velocity of the trajectory is simply the raising of the momentum, that is

$$V := d\phi^\sharp = -\frac{1}{s\sqrt{4s^2 - 1}} \frac{\partial}{\partial s} - \frac{1}{s^2} \frac{\partial}{\partial \theta}.$$

Note that $g(V, V) = 0$. Figure (4.2) shows a plot of the vector field V overlaid on top of the contour surfaces of ϕ for $k = 2$.

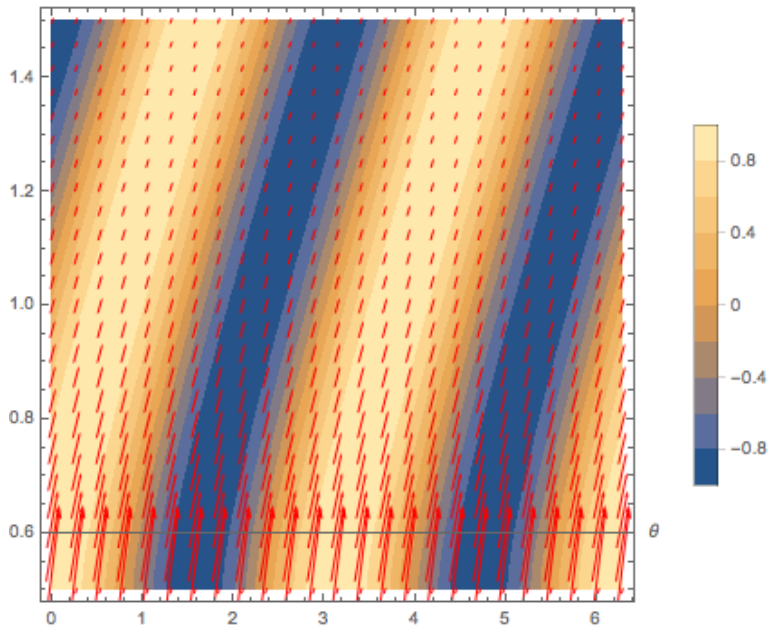


Figure 4.2: Contour plot of the phase function W for $k = 2$, with gradient vector field overlaid on top.

The integral curves of this vector field V are determined by the following equations for a path $(\gamma^s(\lambda), \gamma^\theta(\lambda))$:

$$\frac{d\gamma^s}{d\lambda} = -\frac{1}{\gamma^s\sqrt{4(\gamma^s)^2 - 1}} \qquad \frac{d\gamma^\theta}{d\lambda} = -\frac{1}{(\gamma^s)^2}.$$

Or alternatively

$$\frac{d\gamma^s}{d\lambda}\gamma^s\sqrt{4(\gamma^s)^2-1}=-1 \qquad \frac{d\gamma^\theta}{d\lambda}(\gamma^s)^2=-1.$$

Taking an extra λ derivative of each of the equations, and then isolating the highest derivative leads

$$\frac{d^2\gamma^s}{d\lambda^2}+\frac{4\gamma^s}{4(\gamma^s)^2-1}\left(\frac{d\gamma^s}{d\lambda}\right)^2+\frac{1}{\gamma^s}\left(\frac{d\gamma^s}{d\lambda}\right)^2=0$$

$$\frac{d^2\gamma^\theta}{d\lambda^2}+\frac{2}{\gamma^s}\frac{d\gamma^s}{d\lambda}\frac{d\gamma^\theta}{d\lambda}=0$$

The third term in the first equation can be rewritten, using the original integral curves equations, as follows

$$\frac{1}{\gamma^s}\left(\frac{d\gamma^s}{d\lambda}\right)^2=\frac{1}{(\gamma^s)^3(4(\gamma^s)^2-1)}=\frac{\gamma^s}{4(\gamma^s)^2-1}\left(\frac{d\gamma^\theta}{d\lambda}\right)^2,$$

which when plugged back in, results in

$$\frac{d^2\gamma^s}{d\lambda^2}+\frac{4\gamma^s}{4(\gamma^s)^2-1}\left(\frac{d\gamma^s}{d\lambda}\right)^2+\frac{\gamma^s}{4(\gamma^s)^2-1}\left(\frac{d\gamma^\theta}{d\lambda}\right)^2=0$$

$$\frac{d^2\gamma^\theta}{d\lambda^2}+\frac{2}{\gamma^s}\frac{d\gamma^s}{d\lambda}\frac{d\gamma^\theta}{d\lambda}=0.$$

These are precisely the affinely parameterized geodesic equation, as was to be shown.

4.2 Investigating $\mu = 0$ Solutions

So far, the context of these investigations has been that of a real scalar field Ψ on a real manifold. The use of complex exponential solutions to the wave equation is merely a matter of convenience, standing in place of the oscillatory functions $\sin(x)$ and $\cos(x)$. This is a common approach when dealing with wave equations, since complex exponentials are much easier to deal with. Ultimately, so long as the operations performed on the solution are linear over complex addition (addition, derivation, real scalar multiplication, etc.), keeping only the real part or the imaginary part at the end of the manipulations is like having dealt with only real functions from the beginning. However, this is certainly not the case when building a separation of variables solution

of the form $\psi = S(s)\Theta(\theta)$, where $S(s)$ and $\Theta(\theta)$ have non-zero imaginary parts. In that case,

$$\Re(\psi) \neq \Re(S(s))\Re(\Theta(\theta)).$$

But while there is mixing of real and imaginary parts in the above situation, it still leads to a function whose real and imaginary parts are independently solutions of the original equation. In fact it leads to a basis of solutions that can then be used to construct any generic solution via linear combinations. So again, the using of complex exponentials is purely a convenient way to manipulate an expression, which will ultimately only to be used for either its real or its imaginary parts, both being real valued functions.

The solutions to the Klein-Gordon equation for the case of $\mu = 0$, obtained through the separation of variables method, are

$$\left\{ e^{ik\left(\sqrt{4s^2-1}-\tan^{-1}\left(\sqrt{4s^2-1}\right)\pm\theta\right)} \right\}_{k\in\mathbb{Z}\setminus\{0\}}$$

and linear combinations thereof. Writing these in terms of real valued functions, for $s > 1/2$ yields

$$\left\{ \begin{array}{c} \cos\left(k\left(\sqrt{4s^2-1}-\tan^{-1}\left(\sqrt{4s^2-1}\right)\pm\theta\right)\right), \\ \sin\left(k\left(\sqrt{4s^2-1}-\tan^{-1}\left(\sqrt{4s^2-1}\right)\pm\theta\right)\right) \end{array} \right\}_{k\in\mathbb{N}}$$

where now k need only be in the natural numbers to enumerate the entire basis. This basis of solutions represent left and right moving waves with phase velocity

$$\frac{\sqrt{4s^2-1}}{s}.$$

Figure (4.3) below shows the the right moving cosine solution with $k = 2$, for 6 different values of s.

On the other hand, for for $s < 1/2$, the radical $\sqrt{4s^2-1}$ becomes imaginary, and the sine and cosine solutions becomes complex. Turning back to the solutions in terms of complex exponentials, they can be rewritten as

$$\left\{ e^{k\left(\sqrt{1-4s^2}-\tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)} e^{\pm ik\theta} \right\}_{k\in\mathbb{Z}\setminus\{0\}},$$

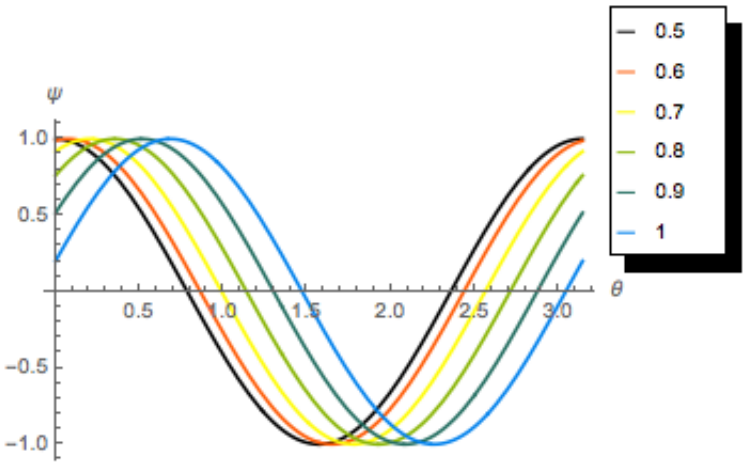


Figure 4.3: Plot of right moving cosine solution with wavenumber $k = 2$ for the s-values $\{0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$.

where the identity $\tan^{-1}(ix) = i\tanh^{-1}(x)$ was used. Furthermore, these can be re-expressed in terms of real valued functions:

$$\left\{ e^{\pm k\left(\sqrt{1-4s^2} - \tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)} \cos(k\theta), \right.$$
$$\left. e^{\pm k\left(\sqrt{1-4s^2} - \tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)} \sin(k\theta) \right\}_{k\in\mathbb{N}}.$$

These solutions are oscillatory in the θ direction, but growing or decaying in the s direction. Although it appears that the growth/decay is exponential in s, using the identity

$$\tanh^{-1}(x) = \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right),$$

the amplitude factor can be rewritten as

$$e^{\pm k\sqrt{1-4s^2}} \left(\frac{2s}{1+\sqrt{1-4s^2}}\right)^{\pm k}$$

which is dominated by the polynomial term. The $+k$ solutions start with zero amplitude at $s = 0$ and grow to max amplitude of 1 at $s = 1/2$, while the $-k$ solutions start with infinite amplitude at $s = 0$ and decay to min overall amplitude of 1 at $s = 1/2$. Figure (4.4) below shows the growing and decaying solutions with wavenumber $k = 2$ in the Riemannian Domain, for 5 different values of s each.

This situation is reminiscent of tunneling problems in Quantum mechanics; at least the mathematics of it, if not the physical interpretation. For a single particle

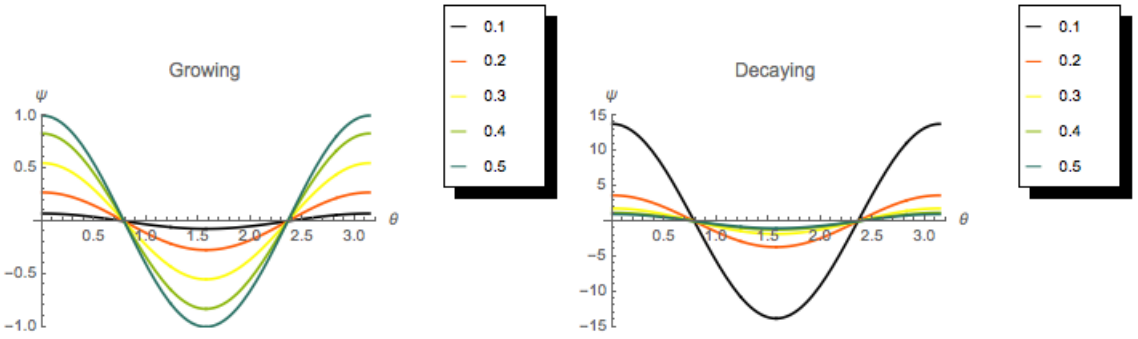


Figure 4.4: Plot of growing and decaying cosine solution with wavenumber $k = 2$ for the s-values $\{0.1, 0.2, 0.3, 0.4, 0.5\}$.

in one dimension, subject to a potential $V(x)$, the time independent Schrödinger equation for the wavefunction Ψ takes the form

$$\frac{d^2\Psi}{dx^2} = -\frac{2m(E - V(x))}{\hbar^2}\Psi(x).$$

For those x such that $E > V(x)$, the equation is approximately of the form

$$\Psi'' = -k^2\Psi,$$

whose solutions behave like traveling waves. On the other hand for the x such that $E < V(x)$, the equation is approximately of the form

$$\Psi'' = \kappa^2\Psi,$$

whose solutions behave like growing or decaying exponentials, also called evanescent waves. Indeed this similarity will be taken further, when applying one of the most powerful techniques of analysis in tunneling problems to the situation at hand, namely the WKB approximation (see chapter 9).

4.3 Preliminary Matching

As a first step towards understanding the behavior of solutions at the transition, a well-behaved, global solution is sought. To this end I impose the condition that ψ be regular at $s = 0$. This is effectively the condition imposed the the Hawking-Hartle no boundary proposal, applied to the situation at hand. This simply requires that

the solution in the Riemannian domain ($s \in (0, 1/2)$) contain none of the negative exponential solutions. So, it must be a linear combination of

$$\left\{ e^{k\left(\sqrt{1-4s^2}-\tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)} \cos(k\theta), \right.$$
$$\left. e^{k\left(\sqrt{1-4s^2}-\tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)} \sin(k\theta) \right\}_{k\in\mathbb{N}}.$$

or alternatively, it must be of the form

$$\sum_{k\in\mathbb{N}} e^{k\left(\sqrt{1-4s^2}-\tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)} \left(\alpha_k \cos(k\theta) + \beta_k \sin(k\theta)\right)$$

for arbitrary real constants α_k, β_k for $k \in \mathbb{N}$. For simplicity, consider a generic, single k cosine wave solution

$$\psi_R(s,\theta) = e^{k\left(\sqrt{1-4s^2}-\tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)} \cos(k\theta).$$

To extend this solution to a unique and differentiable solution on the entire s domain, a solution on the lorenztian domain is sought, such that the value and first derivative at $s = 1/2$ match the value and first derivative of the solution on the Riemannian side, over the entire θ range. In particular, the general solution on the Lorentzian domain can be written as a sum of the form

$$\psi_L(s,\theta) = \sum_{k\in\mathbb{N}} \alpha_k \cos\left(k\left(f(s)+\theta\right)\right) + \beta_k \sin\left(k\left(f(s)+\theta\right)\right) + \tilde{\alpha}_k \cos\left(k\left(f(s)-\theta\right)\right) + \tilde{\beta}_k \sin\left(k\left(f(s)-\theta\right)\right)$$

where

$$f(s) = \sqrt{4s^2-1} - \tan^{-1}\left(\sqrt{4s^2-1}\right),$$

and $\alpha_k, \beta_k, \tilde{\alpha}_k, \tilde{\beta}_k$ for $k \in \mathbb{N}$ are all real numbers. The value of these constants is determined by the matching conditions

$$\psi_R|_{\left(\frac{1}{2},\theta\right)} = \psi_L|_{\left(\frac{1}{2},\theta\right)}, \qquad \left.\frac{\partial\psi_R}{\partial s}\right|_{\left(\frac{1}{2},\theta\right)} = \left.\frac{\partial\psi_L}{\partial s}\right|_{\left(\frac{1}{2},\theta\right)}.$$

Plugging in the expressions for ψ_R and ψ_L into the first of the matching conditions results in

$$\cos(k\theta) = \sum_{k'\in\mathbb{N}} (\alpha_{k'} + \tilde{\alpha}_{k'}) \cos(k'\theta) + (\beta_{k'} + \tilde{\beta}_{k'}) \sin(k'\theta).$$

Using the typical orthogonality of Fourier modes, this equation yields the following conditions

$$\alpha_{k'} + \tilde{\alpha}_{k'} = \delta_{kk'}, \qquad \beta_{k'} + \tilde{\beta}_{k'} = 0.$$

Plugging the expressions for ψ_R and ψ_L into the second matching condition yields $0 = 0$, so no new constraints. This means that the solution on the Lorentzian piece is underdetermined. So, simple continuity and differentiability requirements at the transition boundary are not enough to determine the behavior of the solution across the transition boundary.

The situation becomes a little clearer when considering only the behavior of the S part of the solution. The equation governing S (for $\mu = 0$) is

$$S'' - \frac{1}{s(4s^2-1)}S' + \frac{k^2(4s^2-1)}{s^2}S = 0.$$

A set of linearly independent, real-valued solutions to this equation are

$$S_{L1} = \cos\left(k\sqrt{4s^2-1} - k\tan^{-1}\left(\sqrt{4s^2-1}\right)\right), S_{L2} = \sin\left(k\sqrt{4s^2-1} - k\tan^{-1}\left(\sqrt{4s^2-1}\right)\right)$$

in the Lorentzian domain ($s > 1/2$), and

$$S_{R1} = e^{k\left(\sqrt{1-4s^2} - \tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)}, S_{R2} = e^{-k\left(\sqrt{1-4s^2} - \tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)}.$$

in the Riemanian domain ($0 < s < 1/2$). The values of the solutions and their first derivatives at $s = 1/2$ are

$$\begin{aligned} S_{R1}(1/2) &= 1, & S'_{R1}(1/2) &= 0 \\ S_{R2}(1/2) &= 1, & S'_{R2}(1/2) &= 0 \\ S_{L1}(1/2) &= 1, & S'_{L1}(1/2) &= 0 \\ S_{L2}(1/2) &= 0, & S'_{L2}(1/2) &= 0. \end{aligned}$$

This implies that given any solution $S_R(s) = \alpha S_{R1}(s) + \beta S_{R2}(s)$ in the Riemannian domain and a solution $S_L(s) = \gamma S_{L1}(s) + \delta S_{L2}(s)$ in the Lorentzian side, the matching conditions at $s = 1/2$ impose the single equation

$$\alpha + \beta = \gamma.$$

So, if α, β on the Riemannian side are fixed (either by fiat, or by boundary conditions), then δ is left undetermined, and alternatively if γ, δ are fixed, then α and β are underdetermined. The same conclusion can be reached by considerations of the Wronskians in each of the domains:

$$W[S_{R1}(s), S_{R2}(s)] = -2k\frac{\sqrt{1-4s^2}}{s} \qquad W[S_{L1}(s), S_{L2}(s)] = k\frac{\sqrt{4s^2-1}}{s}.$$

At $s = 1/2$ the Wronskian vanishes, which means that any problem with boundary conditions defined at $s = 1/2$ is not well-posed (see Bender and Orszag [1999]). The matching requirements are effectively a boundary condition problem. However, not all is lost. A way forward can be found through a careful analysis of the local behavior of the solutions. This is the subject of the following chapter.

4.4 Aside: Coordinate Transformation

In section 4.1, the transformation

$$u = \sqrt{4s^2 - 1} - \tan^{-1}\left(\sqrt{4s^2 - 1}\right) \tag{4.3}$$

was used. It was not meant to be taken seriously as a coordinate transformation, rather it was only a little scaffolding used when building the solution to the S equation. However, taking it just a little more seriously it presents an interesting situation. First, note that if the transformation is used as a coordinate transformation, the metric for the toy model becomes

$$f^2 s(u)^2 \left(-du \otimes du + d\theta \otimes d\theta\right) \tag{4.4}$$

which on the face of it looks conformal Minkowski. Even the conformal factor out front is relatively well behaved: it is positive and it only goes to zero at a single point, which corresponds the "south-pole" point of the paraboloid. There is one very important difference however: the coordinate u is not always real. In fact it goes from negative imaginary infinity (when $s = 0$) to zero (when $s = 1/2$) along the negative imaginary axis, and then from zero to positive infinity (as s goes to infinity) along the positive real axis. Geometrically, this has clearly departed the realm of real differentiable manifolds and standard General Relativity.

Interestingly however, it seems that by suspending disbelief and accepting this unusual u behavior, the matching problem has been solved. Indeed, the solution to the temporal part of the separated Klein-Gordon equation was already obtained, at least for the case of $\lambda = 0$. See the paragraphs leading up to equation 4.2. They are

$$e^{iku}, e^{-iku} \tag{4.5}$$

and linear combinations. Furthermore, the Wronskian for these two solutions is

$$W[e^{-iku}, e^{-iku}] = -2ik \tag{4.6}$$

which is nonzero over the entire u domain. In particular, the solutions can be matched trivially at the point $u = 0$. Whatever linear combination of the solutions is used when u is imaginary (the Riemannian domain), matching across $u = 0$ simply means taking the same linear combination when $u > 0$ (the Lorentzian domain). However, this approach is not very rigorous and it raises more questions than it answers, which would have to be analyzed in a broader context like that of complex manifolds and extensions of General Relativity into that domain.

Chapter Five: Interlude - Frobenius Method

This chapter presents a quick overview of the Frobenius method, used to obtain local expansions about regular singular points of linear, ordinary, differential equations.

5.1 Local Solutions

A general n-th order, linear, homogeneous, ordinary differential equation can always be put in the form

$$\frac{d^n y}{dx^n} + p_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}} + .. + p_1(x)\frac{dy}{dx} + p_0(x)y = 0.$$

If the coefficient functions $p_0(x), p_1(x), ..., p_{n-1}(x)$ are continuous over an open interval I, then there are always n linearly independent solutions $y_1(x), ...y_n(x)$ over the interval I. Often, these solutions are not expressible in terms of well known functions. However, a series expression for the solutions might be good enough. In that case, given a point x_0, if the coefficient functions are all analytic in a neighborhood of that point in the complex plane, then all of n linearly independent solutions to the equation are themselves analytic about that point. Furthermore, if they are expressed as a Taylor series

$$y(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k,$$

then their radius of convergence is at least as large as the distance to the closest singular point of the coefficient functions, in the complex plane. Such a point x_0 is called an **ordinary point**. On the other hand, if the coefficient functions are not analytic about x_0, then it is called a **singular point**.

It turns out, however, that it is often the singular points that are of interest. For example, the in transition universe considered in Ch 4, the equation governing the temporal portion $S(s)$ of a separable Klein-Gordon solution takes the form

$$S'' - \frac{1}{s(4s^2-1)}S' + \frac{(k^2 + \lambda^2 s^2)(4s^2-1)}{s^2}S = 0.$$

The main point of interest is $s = 1/2$, where the transition between the Riemannian domain and the Lorentzian domain occurs. This is a singular point of the equation,

so there is no guarantee that the solutions to this equation will be expressible in terms of converging Taylor series about that point. However, progress can still be made in the local analysis of solutions at singular points. In fact, singular points are divided into two categories: *regular singular points* and *irregular singular points*. For an equation of the form

$$\frac{d^n y}{dx^n} + p_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}} + .. + p_1(x)\frac{dy}{dx} + p_0(x)y = 0,$$

if while the coefficient functions themselves are singular at x_0, the following functions

$$\{(x-x_0)p_{n-1}(x), (x-x_0)^2 p_{n-2}(x), (x-x_0)^3 p_{n-3}(x), ..., (x-x_0)^n p_0(x)\}$$

are all analytic in a neighborhood around x_0, then x_0 is called a **regular singular point**. It turns out that while solutions to the above differential equation may not be analytic at regular singular points, if it is singular, the singularity is of a very predictable type. In fact, such an equation must have one solution of the form

$$y(x) = (x-x_0)^\alpha A(x)$$

where α is a (potentially complex) constant called the **indicial exponent** and $A(x)$ is function analytic at x_0. Furthermore, the Taylor series expansion of A must have a radius of convergence at least as large as the distance to the closest singular point of the coefficients (in the complex plane). Any other linearly independent solutions of the differential equation (for order $n \geq 2$) will either be of the form above, with different *indicial exponents* or sums thereof, potentially with factors of

$$\ln(x-x_0)$$

raised to some natural number power in front of some of the sums. For a great presentation of these results see Bender and Orszag [1999].

Any point x_0 that is neither ordinary nor a regular singular point is called an **irregular singular point**. The analysis of local behavior about irregular singular points is not as clear-cut when compared to that of regular singular points, however there is a rich set of tools that can be used to extract a surprising amount of information from such problems. Some of these tools will be discussed as needed in the following chapters. However, for the moment note that for the equation governing the behavior of $S(s)$, the point $s = 1/2$ is a regular singular point. The following section details the standard method for obtaining a local series representation of solutions about regular singular points.

5.2 The Frobenius Method

The Frobenius method is an extension of the usual Taylor series solution approach used for ordinary differential equations, about non-singular points. For the purposes of this investigation, it will be sufficient to consider a 2^{nd} order equation. The extension of this method to higher orders is straightforward. Therefore, consider a generic 2^{nd} order, linear, homogeneous, ordinary differential equation

$$\frac{d^2y}{dx^2} + p(x)\frac{dy}{dx} + q(x)y = 0,$$

such that x_0 is a regular singular point. In particular, this means that $(x - x_0)p(x)$ and $(x - x_0)^2 q(x)$ can be expanded in converging Taylor series around x_0. Therefore, the coefficient functions can be expressed as

$$p(x) = (x - x_0)^{-1} \sum_{k=0}^{\infty} p_k (x - x_0)^k$$
$$q(x) = (x - x_0)^{-2} \sum_{k=0}^{\infty} q_k (x - x_0)^k.$$

Now, the next step is to plug in above expansions of the coefficient functions, and the following ansatz solution

$$y(x) = (x - x_0)^{\alpha} \sum_{k=0}^{\infty} a_k (x - x_0)^k,$$

called a **Frobenius series**, into the differential equation. To simplify the expression, let r represent $(x - x_0)$. The differential equation then becomes

$$\begin{aligned}
&\alpha(\alpha-1)r^{\alpha-2}\sum_{k=0}^{\infty} a_k r^k + 2\alpha r^{\alpha-1}\sum_{k=0}^{\infty} k a_k r^{k-1} + r^{\alpha}\sum_{k=0}^{\infty} k(k-1)a_k r^{k-2} \\
&+ \alpha r^{\alpha-2}\left(\sum_{j=0}^{\infty} p_j r^j\right)\left(\sum_{k=0}^{\infty} a_k r^k\right) + r^{\alpha-1}\left(\sum_{j=0}^{\infty} p_j r^j\right)\left(\sum_{k=0}^{\infty} k a_k r^{k-1}\right) \\
&+ r^{\alpha-2}\left(\sum_{j=0}^{\infty} q_j r^j\right)\left(\sum_{k=0}^{\infty} a_k r^k\right) = 0.
\end{aligned}$$

After re-expressing the multiplied series, reindexing, and collecting by powers of r, the above expression becomes

$$\left(\alpha^2+(p_0-1)\alpha+q_0\right)a_0r^{\alpha-2}+\Big(\left((\alpha+1)^2+(p_0-1)(\alpha+1)+q_0\right)a_1+\left(\alpha p_1+q_1\right)a_0\Big)r^{\alpha-1}$$
$$+\sum_{k=2}^{\infty}\Big(\left((\alpha+k)^2+(p_0-1)(\alpha+k)+q_0\right)a_k+\sum_{j=0}^{k-1}\left((\alpha+j)p_{k-j}+q_{k-j})a_j\right)\Big)r^{k+\alpha-2}=0.$$

Setting the coefficients of each power of r to zero leads to the following equations

$$\left(\alpha^2+(p_0-1)\alpha+q_0\right)a_0=0$$
$$\left((\alpha+k)^2+(p_0-1)(\alpha+k)+q_0\right)a_k+\sum_{j=0}^{k-1}\left((\alpha+j)p_{k-j}+q_{k-j})a_j\right)=0$$
$$(\text{ for } k=1,2,..).$$

At this point it is important to realize that there is an ambiguity in the ansatz series

$$y=r^{\alpha}\sum_{k=0}^{\infty}a_kr^k.$$

Indeed, start with a series with any α, say

$$y=r^{3/2}\left(1+r+r^2+r^3+...\right).$$

Here $\alpha=3/2$, $a_0=a_1=...=1$. This series can equivalently be written as

$$y=r^{1/2}\left(r+r^2+r^3...\right)$$

or

$$y=r^{-1/2}\left(r^2+r^3+r^4+...\right),$$

etc., where $\alpha=1/2$ and $-1/2$ respectively. However, for both of these choices of α, the constant a_0 is zero. To settle on a unique series representation, the convention that $a_0\neq 0$ is chosen. This choice can always be made for any given Frobenius series, and as mentioned, it leads to a unique indicial exponent α.

Using said convention, where $a_0\neq 0$, then the first of the equations that came out of the search for a series solution to the differential equation, can be written as

$$\alpha^2+(p_0-1)\alpha+q_0=0.$$

This is called the **indicial equation**, for 2nd order, linear, homogeneous, ordinary differential equations. The solutions to this equation determine the indicial exponent α and the form of the series solution to the differential equation. It should come as no surprise that it is quadratic, and therefore has two solutions, since the differential equation is 2nd order. Now, having determined an α, the coefficients $\{a_n\}$ are determined using the recursion relation derived above, rewritten as

$$\left((\alpha+k)^2+(p_0-1)(\alpha+k)+q_0\right)a_k=-\sum_{j=0}^{k-1}\left((\alpha+j)p_{k-j}+q_{k-j}\right)a_j,$$

for $k\in\mathbb{N}$. However, notice that in order to solve the recursion relation for a_k in terms of the a_j where $j<k$, the factor

$$(\alpha+k)^2+(p_0-1)(\alpha+k)+q_0$$

must be nonzero. Note that this is just the indicial equation with α replaced by $\alpha+k$. This allows the analysis to be separated into different scenarios, depending on the behavior of the solutions of the indicial equation.

Let the solutions of the indicial equations be denoted α_1 and α_2. Without loss of generality, let the real part of α_1 be greater than or equal than the real part of α_2:

$$\Re(\alpha_1)\geq\Re(\alpha_2).$$

Otherwise just reverse the labels. Now, the Frobenius series associated with the indicial exponent α_1 is completely determined by the recursion relation. That is, the factor

$$(\alpha_1+k)^2+(p_0-1)(\alpha_1+k)+q_0$$

cannot be zero, for any $k\in\mathbb{N}$, since the only other root of the equation has real part that is smaller than or equal to α_1, and thus can never be reached by adding a positive integer to α_1. Therefore all the coefficients are determined by the recursion relation, and furthermore it can be shown that the series is convergent in a radius at least as large as the distance to the nearest singularity. So there is always at least one solution in the form of a Frobenius series.

The form of the other linearly independent solution depends on whether the difference between the roots of the indicial equation is an integer or not. If $\alpha_1-\alpha_2\neq\mathbb{N}$, then it is also true that

$$(\alpha_2+k)^2+(p_0-1)(\alpha_2+k)+q_0$$

will not be zero for any $k \in \mathbb{N}$. So the second solutions will also be a converging Frobenius series, with a different indicial exponent. On the other hand, if $\alpha_1 - \alpha_2 = 0, 1, 2, 3..$, then more care must be taken to construct the second solution. This situation is broken down into three possibilities, which yield different forms of the solution:

1. $\alpha_1 = \alpha_2$
2. $\alpha_1 = \alpha_2 + n$ for a positive integer n and the RHS of the recursion relation is not zero for $k = n$,
3. and $\alpha_1 = \alpha_2 + n$ for a positive integer n and the RHS of the recursion relation is equal to zero for $k = n$.

For both case 1 and 2, the second solution will not be a Frobenius series, rather it will be a sum of one Frobenius series with a logarithm times another Frobenius series. This is analogous to the solution of an Euler equation (a.k.a. equidimensional equation) when the characteristic polynomial has repeated roots. Case 3 on the other hand, does lead to a second solution in the form of a Frobenius series. The constant a_n will simply be arbitrary. For a more complete analysis of the Frobenius method, including a detailed construction of the second solution for the three cases above, see Bender and Orszag [1999].

Chapter Six: Klein-Gordon: Approach 2

This chapter details a resolution to the difficulty of matching solutions to the Klein-Gordon equation at the transition surface, encountered in chapter 4, and investigates some of its consequences. It does so through the use of local analysis of ordinary differential equations - namely the Frobenius method. The analytic solutions on either side of the transition are separately matched to a local approximation to a solution that is well defined in a neighborhood of $s = 1/2$.

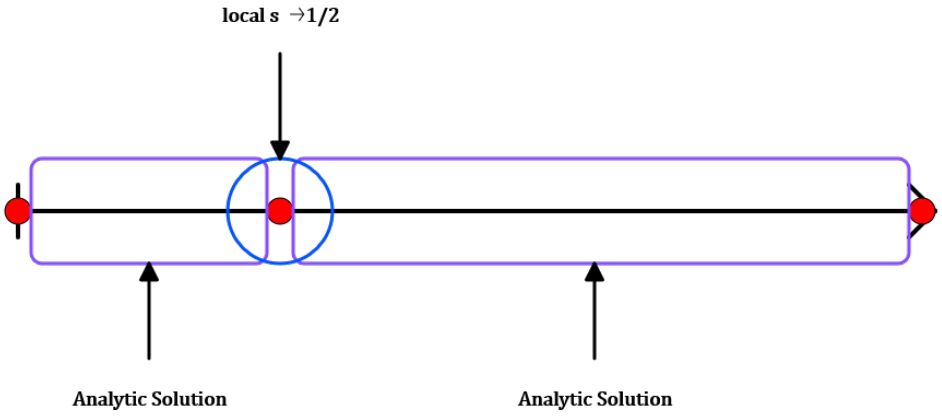
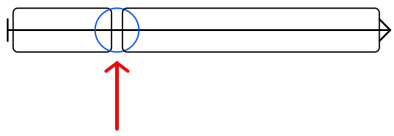


Figure 6.1: These are the relevant domains for the matching approach in this chapter.

6.1 Frobenius expansion about $s = 1/2$



Having seen how the usual continuity and differentiability requirements do not lead to a well-posed problem when attempting to construct unique solutions to the

transition-universe Klein-Gordon equation, the idea now is to use local analysis of solutions about the transition point to gain a better understanding of their behavior across the transition. These local expressions of the solutions will then be matched to the exact solutions obtained for the Riemannian and Lorentzian domains, thereby making the problem well-posed (given appropriate boundary conditions). Therefore I start with a local analysis of the equation governing the evolution of S, about the point $s = 1/2$.

The equation governing the evolution of S, for the case of $\mu = 0$, can be written as

$$S'' + \frac{p(s)}{(s-\frac{1}{2})}S' + \frac{q(s)}{(s-\frac{1}{2})^2}S = 0$$

where

$$p(s) = -\frac{1}{4s(s+\frac{1}{2})} \qquad q(s) = \frac{4k^2(s+\frac{1}{2})(s-\frac{1}{2})^3}{s^2}.$$

The point $s = 1/2$ is a regular, singular point of the equation, since $p(s)$ and $q(s)$ are analytic about that point. The indicial equation takes the form

$$\alpha^2 - \frac{3}{2}\alpha = 0,$$

whose solutions are $\alpha_1 = 3/2$ and $\alpha_2 = 0$. Since the difference between the indicial exponents is neither zero nor a positive integer, both local solutions will be in the form of a Frobenius series. In fact, since $\alpha_2 = 0$, the local solution associated with that indicial exponent will be a Taylor series. Furthermore, neither of the solution will blow up at the transition point.

Constructing the solutions is just a matter of plugging in the coefficients of the Taylor expansions to $p(s)$ and $q(s)$ into the recursion relation

$$\left((\alpha+i)^2 + (p_0-1)(\alpha+i) + q_0\right) a_i = -\sum_{j=0}^{i-1} \left((\alpha+j)p_{i-j} + q_{i-j}\right) \alpha_j.$$

For example, to get the first four terms of either Frobenius series solution, the first four terms in the Taylor expansions of $p(s)$ and $q(s)$ are needed. These are

$$p(s) = -\frac{1}{2} + \frac{3}{2}\left(s-\frac{1}{2}\right) - \frac{7}{2}\left(s-\frac{1}{2}\right)^2 + \frac{15}{2}\left(s-\frac{1}{2}\right)^3 + O(4)$$

$$q(s) = 16k^2\left(s-\frac{1}{2}\right)^3 + O(4)$$

so

$$p_0 = -\frac{1}{2}, p_1 = \frac{3}{2}, p_2 = -\frac{7}{2}, p_3 = \frac{15}{2} \quad \text{and} \quad q_0 = q_1 = q_2 = 0, q_3 = 16k^2.$$

Using these coefficients, the first few terms of the series solutions can be computed. For the indicial exponent $\alpha = 3/2$, the series representation of the solution is

$$y_{3/2}(s) = \left(s - \frac{1}{2}\right)^{3/2} \left(1 - \frac{9}{10}\left(s - \frac{1}{2}\right) + \frac{69}{56}\left(s - \frac{1}{2}\right)^2 - \frac{2}{27}\left(\frac{819}{32} + 16k^2\right)\left(s - \frac{1}{2}\right)^3 + O(4)\right),$$

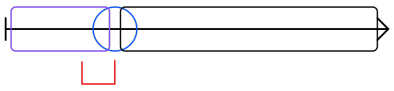
up to an overall, arbitrary constant a_0. While for the indicial exponent $\alpha = 0$ the series is (also up to an overall constant)

$$y_0(s) = 1 - \frac{32k^2}{9}\left(s - \frac{1}{2}\right)^3 + O(4).$$

What these results say is that, in some neighborhood of the points $s = 1/2$, the differential equation for S has two well-defined, non-singular, linearly-independent solutions $y_{3/2}(s)$ and $y_0(s)$, and their local expansions about $s = 1/2$ are the series written above. So any solution to that equation defined on a neighborhood of $s = 1/2$ must be a linear combination of $y_{3/2}(s)$ and $y_0(s)$, and will therefore have a local series solution that is a linear combination of of the series representations of $y_{3/2}(s)$ and $y_0(s)$. This information is sufficient to understand how solutions transition the point $s = 1/2$, or in other words to turn a well posed problem in one of the domains, into a well posed problem over the entire s interval.

6.2 Matching Using Local Expansion

Matching Riemannian Solutions



To accomplish this goal, the next step is to expand the exact, linearly-independent, real solutions on both the Riemannian and the Lorenztian domains, about the point $s = 1/2$. This can be done simply by writing $s = 1/2 \pm \epsilon$ for very small ϵ, with the

sign chosen appropriately depending on the domain considered. For example, the two solutions considered in the Riemannian domain are

$$S_{R1}(s) = e^{k\left(\sqrt{1-4s^2}-\tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)}$$
$$S_{R2}(s) = e^{-k\left(\sqrt{1-4s^2}-\tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)}.$$

Writing $s = 1/2 - \epsilon$ and expanding in orders of ϵ yields

$$S_{R1} = 1 - \frac{8k}{3}\epsilon^{3/2} - \frac{12k}{5}\epsilon^{5/2} + \frac{32k^2}{9}\epsilon^3 - \frac{23k}{7}\epsilon^{7/2} + O(4)$$
$$S_{R2} = 1 + \frac{8k}{3}\epsilon^{3/2} + \frac{12k}{5}\epsilon^{5/2} + \frac{32k^2}{9}\epsilon^3 + \frac{23k}{7}\epsilon^{7/2} + O(4).$$

On the other hand, writing the local solutions $y_{3/2}(s)$ and $y_0(s)$ in terms of $\epsilon = 1/2 - s$ gives

$$y_{3/2} = i\left(\epsilon^{3/2} + \frac{9}{10}\epsilon^{5/2} + \frac{69}{56}\epsilon^{7/2} + \frac{2}{27}\left(\frac{819}{32} + 16k^2\right)\epsilon^{9/2} + O(11/2)\right)$$
$$y_0 = 1 + \frac{32k^2}{9}\epsilon^3 + O(4).$$

where $(s-1/2)^{1/2}$ was written as $i\epsilon^{1/2}$ in the first equation. Here, the choice of letting $(-\epsilon)^{1/2}$ be equal to $i\epsilon^{1/2}$ was made. The alternative root of unity could have been used instead, namely $(-\epsilon)^{1/2} = -i\epsilon^{1/2}$. Ultimately the physics is independent of that choice. However, the choice must be made consistently throughout the analysis.

Now, notice that the solutions S_{R1} and S_{R2} can be matched to linear combinations of $y_{3/2}$ and y_0:

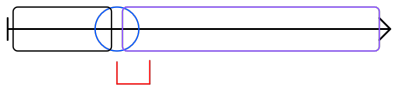
$$S_{R1} = y_0 + i\frac{8k}{3}y_{3/2}, \qquad S_{R2} = y_0 - i\frac{8k}{3}y_{3/2}, \tag{6.1}$$

at least up to $O(4)$. It should come as no surprise that this is the case, since any solution in a domain that overlaps the radius of convergence of the $y_{3/2}$ and y_0, must be expressible as a linear combination of y_0 and $y_{3/2}$. Furthermore, the above relationships will hold exactly, no matter to how many orders the series representations are developed.

These expressions of S_{R1} and S_{R2} in terms of $y_{3/2}$ and y_0 are useful because, while it was unclear how to continue the solutions in the Riemannian domain into the Lorentzian domain, there is a sense in which the solutions $y_{3/2}$ and y_0 naturally transition across $s = 1/2$. Typical existence-uniqueness theorems guarantee that y_0

and $y_{3/2}/(s-1/2)^{3/2}$ are analytic at $s = 1/2$, and the radius of convergence is at least as large as the distance to the nearest other singular point of the coefficients in the complex plane. Therefore, by analytic continuation, the power series expansion on one side of the transition point, determines the functions on the other side. Of course, using the analyticity of these solutions to transition them across the singular point represents a choice, and that is the choice to consider only analytic solutions to the S equation. If this restriction is lifted then there are many piecewise defined solutions to the S equation that are just as good as the $y_{3/2}$ solution but the behavior on one side of $s = 1/2$ does not completely pin down the behavior on the other side. However, I move forward with the restriction to analyticity, for the pragmatic reason that it allows a way forward. For an interesting discussion about the use of analyticity as a fundamental principle in physics see Chew [1965].

Matching Lorentzian Solutions



By similarly expanding the solutions which are real in the Lorentzian domain,

$$S_{L1}(s) = \cos\left(k\sqrt{4s^2-1} - k\tan^{-1}\left(\sqrt{4s^2-1}\right)\right)$$
$$S_{L2}(s) = \sin\left(k\sqrt{4s^2-1} - k\tan^{-1}\left(\sqrt{4s^2-1}\right)\right),$$

close to the transition, they can be identified as linear combinations of $y_{3/2}$ and y_0. These can then be used to write the unique extension of S_{R1} into the Lorentzian domain in terms of S_{L1} and S_{L2}.

To carry this out, let $s = 1/2 + \epsilon$ for a very small but positive ϵ, and expand S_{L1} and S_{L2} in orders of ϵ as follows:

$$S_{L1} = 1 - \frac{32k^2}{9}\epsilon^3 + \frac{32k^2}{5}\epsilon^4 + O(5)$$
$$S_{L2} = \frac{8k}{3}\epsilon^{3/2} - \frac{12k}{5}\epsilon^{5/2} + \frac{23k}{7}\epsilon^{7/2} - \frac{16}{81}k^2\left(\frac{819}{32} + 16k^2\right)\epsilon^{9/2} + O(11/12).$$

Writing $y_{3/2}$ and y_0 in terms of $\epsilon = s - 1/2$ this time, leads to the identification that

$$S_{L1} = y_0 \qquad S_{L2} = \frac{8k}{3}y_{3/2}.$$

Using this information, the matching procedure can be finalized. Starting from any solution in one of the domains, say S_{R1} in the Riemannian domain, it can be extended through the transition by writing it in terms of $y_{3/2}$ and y_0, which in the case of S_{R1} is

$$y_0 + i\frac{8k}{3}y_{3/2}.$$

Then, writing this solution in terms of the real solutions in the new domain, in this case

$$S_{L1} + iS_{L2}.$$

Hence the matching is complete. For the sake of compactness this can be written as

$$S_{R1} \to S_{L1} + iS_{L2},$$

and the analogous relation for S_{R2} is

$$S_{R2} \to S_{L1} - iS_{L2}.$$

This unambiguously fixes the matching problem. The following section discusses some of the main consequences of this matching procedure. However, before moving forward, it is interesting to note that the resolution of the matching dilemma via this Frobenius series approach coincides with the result obtained in section 4.4. This suggests that perhaps taking the complex coordinate transformation a little more seriously is not a bad idea.

6.3 Regularity OR Reality

The first thing to notice from the example above is that, a solution on the Riemannian domain like S_{R1}, which is purely real, transitions into a solution on the Lorentzian domain that is complex:

$$S_{L1} + iS_{L2}.$$

Note that this is fundamentally different than using complex solutions to simplify the algebra and then just keeping the real part. In this case there is no freedom to take linear combinations in the Lorentzian side once the solution on the Riemannian side has been picked (or vice-versa). This was the entire goal of finding a matching procedure. So the complex nature is here to stay. Having said that, if the solutions

S_{R1}, S_{R2}, S_{L1} and S_{L2} are now interpreted as complex valued functions on the entire $s \in (0, \infty)$ domain, then they satisfy the equations

$$S_{R1}(s) = S_{L1}(s) + iS_{L2}(s)$$
$$S_{R2}(s) = S_{L1}(s) - iS_{L2}(s).$$

These are just re-statements of the Euler equation $e^{i\theta} = \cos(\theta) + i\sin(\theta)$, and are also exactly the result of the matching procedure developed above. In other words, the local matching procedure amounts to simply taking each solution at face value, and ignoring the transition point $s = 1/2$. So long as appropriate initial conditions are specified, in particular not at $s = 1/2$, the problem is well-posed and a unique solution exists on the entire s domain. However, the complex nature of the solution must be accepted.

This happens because of the Frobenius solution $y_{3/2}$, whose series representation is

$$\left(s - \frac{1}{2}\right)^{3/2}\left(1 - \frac{9}{10}\left(s - \frac{1}{2}\right) + \frac{69}{56}\left(s - \frac{1}{2}\right)^2 - \frac{2}{27}\left(\frac{819}{32} + 16k^2\right)\left(s - \frac{1}{2}\right)^3 + O(4)\right).$$

This solution, picks up a factor of i as it crosses the point $s = 1/2$. Every solution to the equation can be expressed as a linear combination of $y_{3/2}$ and y_0. So any solution that contains any part of the $y_{3/2}$, cannot be real on both sides of the transition point. So, for example, imposing the condition that the solution be regular at $s = 0$, as distilled from the Hawking-Hartle no boundary condition, implies that the solution on the Riemannian is of the form,

$$S(s) = \alpha S_{R1}(s),$$

where α is an arbitrary constant. This is the case because S_{R2} blows up as s goes to zero. This solution can alternatively be written as

$$S(s) = \alpha\left(S_{L1}(s) + iS_{L2}(s)\right).$$

Given the new understanding obtained from the matching procedure, both expressions are valid over the entire s domain. Note that there is no choice of $\alpha \in \mathbb{C}$ that makes this solution real-valued in the Lorentzian domain. This would require that

$$\Re(\alpha)S_{L2}(s) + \Im(\alpha)S_{L1}(s) = 0.$$

Linearly independence guarantees that the only solution to this condition is $\alpha = 0$.

On the flip side, a solution that is just made out of the y_0 piece can be real throughout. Using the relationships between the y solutions and the S_R, S_L solutions, this means that only solutions that are real valued in the entire s domain cane be written as

$$\beta\left(S_{L1}(s)\right)$$

or, equivalently

$$\beta\left(S_{R1} + S_{R2}\right),$$

where β is a real valued constant. Plugging in the actual form of the solutions, these expressions become

$$\beta\cos\left(k\sqrt{4s^2-1} - k\tan^{-1}\left(\sqrt{4s^2-1}\right)\right), \qquad \beta\cosh\left(k\sqrt{1-4s^2} - k\tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)$$

respectively. Figure (6.2) shows the plot of one choice of real-valued solution through transition, where the overall constant β has been set to 1 and the wavenumber k set to 3. However, as already remarked, this real solution will necessarily blow up at the origin. So the solution can either be regular or real-valued in the entire domain, but not both.

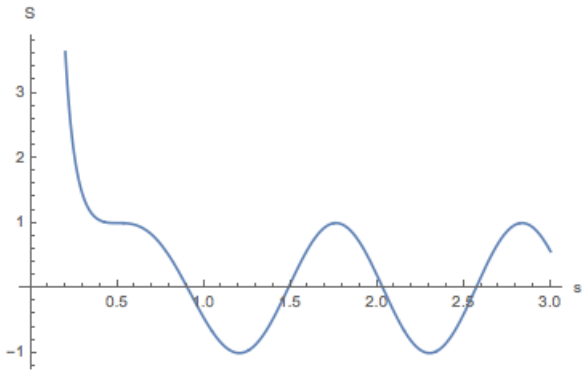


Figure 6.2: Plot of solution to the transition universe Klein-Gordon equation that remains real throughout the entire domain. For concreteness in the plot, the overall constant $\beta = 1$ and the wave number $k = 3$.

There are different possible interpretations of this situation which depend on the physical meaning of the Klein-Gordon equation, and the choice between regularity or reality. For example, if regularity at the origin is chosen, this result can be interpreted as the requirement that a scalar field theory, compatible with a transition

universe, must be charged. A complex Klein-Gordon equation is often interpreted as a pair of real-valued Klein-Gordon fields, with a particular conserved quantity that is interpreted as charge.

6.4 Geodesic limit

Given the new understanding of the transition behavior of the Klein-Grodon solutions, it is instructive to step back and see what information can be obtained about the matching of geodesics through the geometric optics approximation. In section 4.1, the solution

$$\psi(s,\theta) = e^{ik\left(\sqrt{4s^2-1}-\tan^{-1}\left(\sqrt{4s^2-1}\right)-\theta\right)}$$

was used to explore the relationship between Klein-Gordon waves and geodesics. Because of its plane wave form, there was no need to take any limit in order to make the duality manifest. It was shown that calculating the gradient vector field of the phase function

$$k\left(\sqrt{4s^2-1}-\tan^{-1}\left(\sqrt{4s^2-1}\right)-\theta\right)$$

can be integrated (in the sense of integral curves) to obtain particle geodesics. At that time however, the complex exponential solution was being used simply as an algebraic convenience, and ultimately the solution really being considered was

$$\psi(s,\theta) = \cos\left(k\left(\sqrt{4s^2-1}-\tan^{-1}\left(\sqrt{4s^2-1}\right)-\theta\right)\right).$$

With the new light shed on the transition problem through local analysis of the differential equation, more care needs to be taken when discussing complex solutions.

The cosine solution above represents a right moving wave in the Lorentzian domain. It can be decomposed into a sum of separable solutions:

$$\begin{aligned}\psi(s,\theta) = &\cos\left(k\left(\sqrt{4s^2-1}-\tan^{-1}\left(\sqrt{4s^2-1}\right)\right)\right)\cos(k\theta)\\ &+\sin\left(k\left(\sqrt{4s^2-1}-\tan^{-1}\left(\sqrt{4s^2-1}\right)\right)\right)\sin(k\theta)\end{aligned}$$

Therefore this solution contains both S_{L1} and S_{L2} solutions to the S equation, which means that its extension into the Riemannian domain will be both not real and non-regular.

At this point it seems that there are two potential alternatives to consider. First, I can consider a different solution, one which is real throughout the entire domain,

and explore its link to geodesics. A general solution with this property is a sum of standing waves of the form

$$\psi(s,\theta) = \cos\left(k\left(\sqrt{4s^2-1} - \tan^{-1}\left(\sqrt{4s^2-1}\right)\right)\right)\left(a_k\cos(k\theta) + b_k\sin(k\theta)\right).$$

Lamentably, for this solution, the connection to rays and particles has been lost. This is not approximately a plane wave, and there is no such thing as a phase function. So the behavior of this solution across the transition has no bearing on the behavior of geodesic paths across the transition.

The second alternative is to continue working with a solution like

$$\psi(s,\theta) = \cos\left(k\left(\sqrt{4s^2-1} - \tan^{-1}\left(\sqrt{4s^2-1}\right) - \theta\right)\right),$$

keeping in mind that it becomes complex in the Riemannian domain and therefore the connection to waves and particles is tenuous at best. It was already shown, in section 4.1, that this solution corresponds to a family of null geodesics in the Lorentzian domain. The extension of this solution to the Riemannian domain, using the expression of the solution in terms of separable solutions, is

$$\begin{aligned}\psi(s,\theta) =& \frac{1}{2}\left(e^{+k\left(\sqrt{1-4s^2}-\tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)} + e^{-k\left(\sqrt{1-4s^2}-\tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)}\right)\cos(k\theta)\\ &- \frac{i}{2}\left(e^{+k\left(\sqrt{1-4s^2}-\tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)} - e^{-k\left(\sqrt{1-4s^2}-\tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)}\right)\sin(k\theta).\end{aligned}$$

While this again doesn't have the plane wave form of the geometric optics limit, this time taking the limit as $k \to \infty$ does matter. Since the expression

$$\sqrt{1-4s^2} - \tanh^{-1}\left(\sqrt{1-4s^2}\right)$$

is less than or equal to zero for $s \in (0, 1/2]$, the dominant terms as $k \to \infty$ are

$$\frac{1}{2}e^{-k\left(\sqrt{1-4s^2}-\tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)}\left(\cos(k\theta) + i\sin(k\theta)\right),$$

which can alternatively be written as

$$\frac{1}{2}e^{ik\left(i\left(\sqrt{1-4s^2}-\tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)+\theta\right)}.$$

This does look somewhat like a plane wave except for the glaring fact that the phase function is complex valued:

$$\phi(s,\theta) = ik\left(\sqrt{1-4s^2} - \tanh^{-1}\left(\sqrt{1-4s^2}\right)\right) + k\theta.$$

Blindly following the optico-mechanical analogy would lead to a complex valued velocity vector field. It is unclear what meaning, if any, can be ascribed to such particles, given that the context has so far been that of a real manifold, with a real tangent space.

So at the moment it seems that the Klein-Gordon analysis at the transition does not have anything useful to say about the behavior of particle paths at the transition, but it is not altogether clear why this is happening. As it turns out, two dimensional spacetimes are quite unique vis-a-vis the geometric optics approximation. Once higher dimensional spacetimes are considered, the reason for this disconnect will become apparent. Furthermore this will have important implications in the contextual interpretation of this transition universe problem. Getting a handle on the behavior of the Klein-Gordon in higher dimensional transition universe (and also for the case of $\mu \neq 0$) will require the use of more advance tools than those that have been used so far. In particular, furthering the similarity between this problem and tunneling problems in quantum mechanics, I use asymptotic analysis and a WKB-like approximation. These techniques are described in the chapter 8. However, before doing that, the generalization of the transition universe model to higher spatial dimensions is described in chapter 7.

Chapter Seven: Higher Dimensional and Massive Generalization

This chapter contains a generalization of the toy model considered so far. Both higher dimensions and non-zero masses are considered. The cosmological nature of the model is discussed, including the energy and pressure content of the universe and the existence of cosmological and particle horizons. A different model is presented which exhibits a similar Riemannian to Lorentzian transition in the early times, but then behaves like a generic (flat) FLRW spacetime. Finally, the Klein-Gordon equation in the general mass, general dimensions is described.

7.1 Generalization to higher dimensions

The generalization of the toy model transition universe to higher dimensions is straight forward. Start with a $\mathbb{R}^{(n+1)}$ with the Minkowski metric expressed in generalized cylindrical polar coordinates $(cT, \rho, \theta_1, ..., \theta_{n-1})$. These coordinates are defined by the parameterization

$$\begin{aligned} cT &= cT \\ X_1 &= \rho\cos(\theta_1) \\ X_2 &= \rho\sin(\theta_1)\cos(\theta_2) \\ &\vdots \\ X_{n-1} &= \rho\sin(\theta_1)\sin(\theta_2)...\sin(\theta_{n-2})\cos(\theta_{n-1}) \\ X_n &= \rho\sin(\theta_1)\sin(\theta_2)...\sin(\theta_{n-2})\sin(\theta_{n-1}) \end{aligned}$$

where cT takes values in $(-\infty, \infty)$, the coordinate ρ takes values in $(0, \infty)$, the coordinates θ_1 through θ_{n-2} take values in $(0, \pi)$, and the coordinate θ_{n-1} takes values in $(0, 2\pi)$. In these coordinates, the Minkowski metric takes the form

$$-d(cT) \otimes d(cT) + d\rho \otimes d\rho + \rho^2 d\Omega_{n-1}^2,$$

where $d\Omega_{n-1}^2$ is the metric on the $n-1$ sphere, which takes the form

$$d\theta_1 \otimes d\theta_1 + \sin^2(\theta_1)d\theta_2 \otimes d\theta_2 + ... + \sin^2(\theta_1)...\sin^2(\theta_{n-2})d\theta_{n-1} \otimes d\theta_{n-1}.$$

Now define the transition universe model as the submanifold defined by the equation

$$(cT) = \frac{\rho^2}{f}.$$

The induced metric is thus

$$g = -\left(\frac{4(ct)^2}{f^2} - 1\right) d(ct) \otimes d(ct) + (ct)^2 d\Omega_{n-1}^2, \tag{7.1}$$

where as before ρ was replaced by ct since it is a timelike coordinate; at least after the transition.

Before moving on to an analysis of the Klein-Gordon equation, it is interesting to consider some of the physics of the toy model. For the sake of concreteness, let $n = 4$. That is, consider a $4+0$ to $3+1$ transition. In this case the metric takes the form

$$-\left(\frac{4(ct)^2}{f^2} - 1\right) d(ct)\otimes d(ct) + (ct)^2 \left(d\chi \otimes d\chi + \sin(\chi)^2 d\theta \otimes d\theta + \sin^2(\chi)\sin^2(\theta) d\phi \otimes d\phi\right) \tag{7.2}$$

where for simplicity, the angles $(\theta_1, \theta_2, \theta_3)$ have been relabeled as (χ, θ, ϕ). Letting $s = (ct)/f$, the nonzero components of the connection coefficients, Rieman tensor, Ricci tensor and Ricci scalar can be written as (no summation implied over any indices)

- Connection coefficients

$$\Gamma^0{}_{00} = \frac{4s}{4s^2-1} \qquad \Gamma^0{}_{ii} = \frac{1}{f^2 s(4s^2-1)} g_{ii} \qquad \Gamma^i{}_{i0} = \Gamma^i{}_{0i} = \frac{1}{s}$$

$$\Gamma^1{}_{22} = -\cos(\chi)\sin(\chi) \quad \Gamma^1{}_{33} = -\cos(\chi)\sin(\chi)\sin^2(\theta) \quad \Gamma^2{}_{33} = -\cos(\theta)\sin(\theta)$$

$$\Gamma^2{}_{21} = \Gamma^2{}_{12} = \frac{\cos(\chi)}{\sin(\chi)} \quad \Gamma^3{}_{31} = \Gamma^3{}_{13} = \frac{\cos(\chi)}{\sin(\chi)} \qquad \Gamma^3{}_{32} = \Gamma^3{}_{23} = \frac{\cos(\theta)}{\sin(\theta)}$$

- Riemann tensor

$$R^0{}_{ii0} = -R^0{}_{i0i} = \frac{4}{f^2(4s^2-1)^2} g_{ii}$$

$$R^i{}_{0i0} = -R^i{}_{00i} = \frac{4}{4s^2-1}$$

$$R^i{}_{jij} = -R^i{}_{jji} = \frac{4}{f^2(4s^2-1)} g_{jj} \quad (j \neq i)$$

- Ricci tensor
$$R_{00} = \frac{12}{4s^2 - 1} \qquad R_{ii} = \frac{4(8s^2 - 3)}{f^2(4s^2 - 1)^2} g_{ii}$$
- Ricci Scalar
$$R = \frac{48(2s^2 - 1)}{f^2(4s^2 - 1)^2}.$$

These are indeed very similar to the 2 dimensional case, with notable the exception being the Ricci scalar. In 2D the Ricci scalar was of the form

$$R = -\frac{8}{f^2(4s^2 - 1)^2},$$

which is negative for all s (the dimensional constant f has been added), and approaches zero from below as s goes to infinity. On the other hand, the Ricci scalar for the 4D case behaves similar to the 2D case in the Riemannian domain, but in the Lorentzian domain it goes from being negative infinity at $s = 1/2$, through zero at $s = 1/\sqrt{2}$, to a maximum value of 6 (in units of f^{-2}) at s equals $\sqrt{3}/2$. Then it approaches zero from above as s goes to infinity.

Furthermore, contrary to the 2D case, the Einstein tensor does not vanish exactly. Instead, it is equal to

$$G_{00} = -\frac{12}{f^2(4s^2 - 1)} g_{00}, \qquad G_{ii} = -\frac{4(4s^2 - 3)}{f^2(4s^2 - 1)^2} g_{ii}.$$

From Einstein's field equation

$$G_{\mu\nu} = \kappa T_{\mu\nu}, \qquad \text{where } \kappa = \frac{8\pi G}{c^4},$$

the stress-energy-momentum tensor associated with this metric takes the form

$$T_{00} = -\frac{12}{\kappa f^2(4s^2 - 1)} g_{00}, \qquad T_{ii} = -\frac{4(4s^2 - 3)}{\kappa f^2(4s^2 - 1)^2} g_{ii}. \tag{7.3}$$

This has the form of the stress-energy-momentum tensor for a homogeneous, isotropic perfect fluid

$$T_{\mu\nu} = (\rho + p) U_\mu U_\nu + p g_{\mu\nu} \tag{7.4}$$

where U is the co-moving four-velocity of the fluid, which is freely falling. To see that equation (7.3) is of the form (7.4), it is important to realize that the vector field with components

$$(1, 0, 0, 0)$$

is not the four-velocity of free-falling particles. This is due to the use of a time coordinate that is not *cosmic time*. On the other hand, since cosmic time τ and the coordinate s are related via the coordinate transformation defined by

$$f\sqrt{4s^2-1}ds = d\tau \tag{7.5}$$

one can obtain the four-velocity field that is tangent to freely falling particles via the coordinate transformation law between the cosmic time and s time coordinates (leaving all the other coordinates as is). This yields

$$U^\mu = \left(\frac{1}{f\sqrt{4s^2-1}}, 0, 0, 0\right). \tag{7.6}$$

The integral curves for this vector field can be shown to satisfy the geodesic equation. From equation (7.6) and the metric, it follows that

$$U_0 U_0 = f^2(4s^2-1) = -g_{00}$$

therefore the stress-energy-momentum tensor (7.3) can be put into the form (7.4) with

$$\rho = \frac{12}{\kappa f^2(4s^2-1)} \qquad p = -\frac{4(4s^2-3)}{\kappa f^2(4s^2-1)^2}. \tag{7.7}$$

Notice that the units of these, supplied by the $1/(\kappa f^2)$ factor, are indeed the units of energy density and pressure. Figure (7.1) is a plot of the energy density and pressure, in units of κf^2. The energy density goes from positive infinity after the transition, down to zero as s goes to infinity. The pressure goes from positive infinity after the transition, through zero, down to a minimum value of $-1/2$ and then back up to zero from below.

The argument leading to the identification of the Einstein tensor to energy density and pressure of a perfect fluid does not hold for the Riemannian domain. First, the coordinate transformation defined by (7.5) would have to be modified, to avoid complex coordinates. Furthermore, the general form for a homogeneous, isotropic perfect fluid stress-energy-momentum tensor (7.4) no longer applies. Consider, for example a space with a euclidean metric, in coordinates where the components fo $g_{\mu\nu}$ are $\delta_{\mu\nu}$ and $U^\mu = (1, 0, 0, 0)$. In that case, $T_{00} = (\rho + 2p)$. However, the more fundamental problems is that our understanding of energy and momentum are intimately tied to the spacetime character of the geometry. So it is not clear what the meaning of the stress-energy-momentum tensor would be in the Riemannian domain.

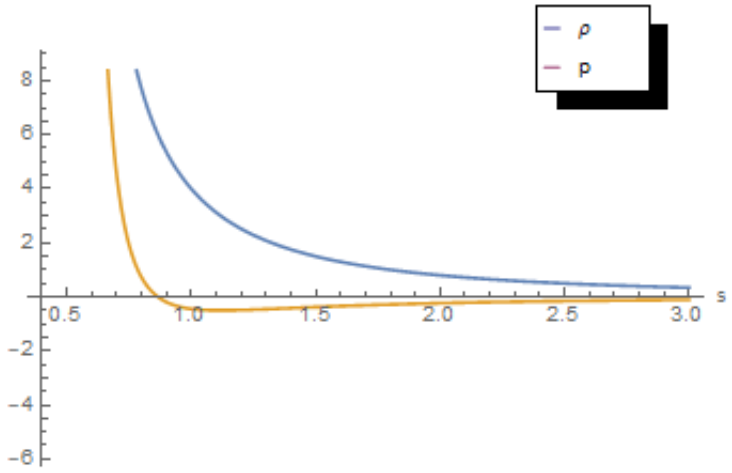


Figure 7.1: Energy density ρ and pressure p in units of κf^2, for 4D transition universe toy model.

7.2 Cosmology

The metric (7.2) can be put into Friedman-Lemâitre-Robertson-Walker (FLRW) form, for $s > 1/2$, by using cosmic time coordinate τ (see paragraph preceding equation 2.1), defined by the equation

$$d\tau = f\sqrt{4s^2 - 1}ds. \tag{7.8}$$

This yields

$$-d\tau \otimes d\tau + (fs(\tau))^2 \left(d\chi \otimes d\chi + \sin(\chi)^2 d\Omega_2^2\right) \tag{7.9}$$

where

$$\frac{\tau}{f} = \frac{1}{2}s\sqrt{4s^2 - 1} - \frac{1}{4}\ln\left(2s + \sqrt{4s^2 - 1}\right). \tag{7.10}$$

While this cannot be solved for s in terms of τ explicitly, it is still interesting to move forward and push the analysis a little further using the implicit relationships between s and τ. The function $fs(\tau)$ is now the scale factor in the typical FLRW analysis.

First, note that the expressions (7.7) for ρ and p in terms of s, together with the relationship

$$f\dot{s} = \frac{1}{\sqrt{4s^2 - 1}}, \tag{7.11}$$

(where the dot refers to derivative with respect to cosmic time) satisfy the continuity and Friedmann equations for a closed universe (curvature constant $k = +1$); namely

$$\begin{aligned} \dot{\rho} &= -3\frac{\dot{s}}{s}(\rho + p) \\ \left(\frac{\dot{s}}{s}\right)^2 + \frac{1}{f^2 s^2} &= \frac{\kappa}{3}\rho \\ \frac{\ddot{s}}{s} &= -\frac{\kappa}{6}(\rho + 3p). \end{aligned} \tag{7.12}$$

using scale factor equal to $fs(\tau)$. As before, $\kappa = 8\pi G/c^4$, not to be confused with the curvature constant k which has already been set to 1. The fact that the Friedmann equations are satisfied comes as no surprise given that they are simply Einstein's field equations for an FLRW metric, and the expressions for ρ and p above were obtained by enforcing precisely Einstein's field equations.

Continuing the analysis, note that the ratio of pressure p to energy density ρ, is not constant. In terms of the scale factor s, it is

$$w(s) := -\frac{(4s^2 - 3)}{3(4s^2 - 1)}. \tag{7.13}$$

This goes from positive infinity at $\tau = 0$ ($s = 1/2$), and approaches $-1/3$ monotonically from above as τ goes to infinity ($s \to \infty$). Both the *weak* and *strong energy conditions* hold since both

$$\begin{aligned} \rho + p &= \frac{32s^2}{\kappa f^2 (4s^2 - 1)^2} \\ \rho + 3p &= \frac{24}{\kappa f^2 (4s^2 - 1)^2} \end{aligned}$$

are greater than zero for $s > 1/2$. As far as the ratio of energy density to critical energy density $\Omega = \rho/\rho^{cr}$, it is

$$\begin{aligned} \Omega &= 1 + \frac{1}{f^2 \dot{s}^2} \\ &= 4s^2. \end{aligned} \tag{7.14}$$

The critical energy density ρ^{cr} is the energy required for the universe to be flat, in terms of the Hubble parameter

$$\rho^{cr} = \frac{3H^2}{\kappa}.$$

Note that $\Omega > 1$ for all $s > 1/2$ as expected from a closed universe and it grows as s grows.

This spacetime has a particle horizon but no event horizon. Radial, null geodesics satisfy

$$\chi = \int_{\tau_i}^{\tau_f} \frac{d\tau}{fs(\tau)}. \tag{7.15}$$

The particle horizon is obtained by letting $\tau_i = 0$ and $\tau_f = \tau_*$, the time of interest.

$$\begin{aligned} \chi_p(\tau_*) &:= \int_0^{\tau_*} \frac{d\tau}{fs(\tau)} \\ &= \int_{1/2}^{s_*} \frac{ds}{f\dot{s}s} \\ &= \int_{1/2}^{s_*} \frac{\sqrt{4s^2-1}}{s} ds \\ &= \sqrt{4s_*^2-1} - \tan^{-1}\left(\sqrt{4s_*^2-1}\right) \end{aligned} \tag{7.16}$$

where $s_* := s(\tau_*)$. Solving the equation $\chi_p(s^*) = \pi$ yields $s = 2.3107...$ For any s less than that value there is a particle horizon. On the other hand the event horizon is obtained by letting $\tau_i = \tau_*$ and $\tau_f = \infty$. That integral diverges so there is no event horizon.

For late times, nothing is particularly out of the ordinary. Playing the movie of the universe in reverse looks like an ordinary big-bang universe until it gets closer and closer to the singularity in energy density and pressure. Perhaps the most salient peculiarity of this spacetime as a cosmological model is the fact that the energy density and pressure blow up before the scale factor reaches zero (when running time in reverse). This is directly linked to the nature of the singularity in this transition universe example. An interesting question is what sort of observational evidence could point to such a "unaligned" singularity, particularly if the universe is opaque before recombination happens. A potential answer to this question is explored in chapter 11.

7.3 Aside: FLRW Like Example

This particular toy model that is being considered is a very specified/prescriptive evolution. It has been chosen as tool to investigate the transition phenomenon for

its simplicity. Even the late time evolution of the model does not reflect our current understanding of the cosmological nature of our universe. However, it is not difficult to obtain spacetimes that exhibit a similar transition singularity at early times and act move like the typical FLRW solutions at later times. For example, consider that following metric on an $\mathbb{R}^4$ manifold:

$$-(s-1)ds \otimes ds + \left(\frac{2}{3}s^{3/2}\right)^q \left(dr \otimes dr + r^2 d\Omega_2^2\right) \tag{7.17}$$

where $s, r \in (0, \infty)$, $d\Omega_2^2$ is the typical metric on a two-sphere and q is an undetermined real number. This metric exhibits a transition at $s = 1$. Now, for $s > 1$, during the Lorentzian phase, cosmic time is defined via the condition

$$\sqrt{s-1}ds = dt.$$

This can be integrated and solved for s in terms of t, which yields

$$s(t) = \frac{3}{2}t^{2/3} + 1.$$

Plugging this back into the metrics (7.17) yields

$$-dt \otimes dt + \left(\frac{2}{3}\left(\left(\frac{3}{2}t\right)^{2/3} + 1\right)^{3/2}\right)^q (dr \otimes dr + r^2 d\Omega_2^2).$$

For $t >> 1$ this metric is approximately

$$-dt \otimes dt + t^q(dr \otimes dr + r^2 d\Omega_2^2)$$

which for different choices of q can model some cosmologies that are typically considered.

7.4 Klein-Gordon Field

Returning to the analysis of a Klein-Gordon field and general dimensions n, the Klein-Gordon action is

$$S_{KG} = \alpha \int \left(-\frac{1}{2}g^{\mu\nu}\partial_\mu\Psi\partial_\nu\Psi - \frac{1}{2}\frac{\mu^2c^2}{\hbar^2}\Psi^2\right)\sqrt{|g|}\, d(ct) \wedge d\theta_1 \wedge ... \wedge d\theta_{n-1},$$

which implies the Klein-Gordon equation

$$\frac{1}{\sqrt{-g}}\partial_\mu\left(\sqrt{-g}g^{\mu\nu}\partial_\nu\Psi\right)=\frac{\mu^2c^2}{\hbar^2}\Psi.$$

For the transition spacetime under consideration (eq 7.1), the KG equation takes the form

$$\frac{1}{\left(\frac{4(ct)^2}{f^2}-1\right)}\left(\frac{1}{c^2}\Psi_{tt}-\frac{(2-n)\frac{4(ct)^2}{f^2}+(n-1)}{ct\left(\frac{4(ct)^2}{f^2}-1\right)}\frac{1}{c}\Psi_t\right)-\frac{1}{(ct)^2}\Delta_{S^{n-1}}\Psi=-\frac{\mu^2c^2}{\hbar^2}\Psi.$$

In the above equation $\Delta_{S^{n-1}}$ is the Laplace-Beltrami operator on S^{n-1} (also known as the $n-1$ spherical Laplace operator). This operator can be written compactly in terms of the components h_{ij} of the metric $d\Omega^2_{n-1}$ on the $n-1$ sphere. In terms of these components, $\Delta_{S^{n-1}}$ is

$$\Delta_{S^{n-1}}\Psi=\frac{1}{\sqrt{\det(h)}}\partial_i\left(\sqrt{\det(h)}h^{ij}\partial_j\Psi\right).$$

For $n=2$, the metric $d\Omega^2_1$ on the circle S^1 is simply $d\theta\otimes d\theta$, which means that

$$\Delta_{S^1}\Psi=\partial_{\theta\theta}\Psi,$$

in agreement with the previous chapters. While for the case $n=4$ the components of the metric $d\Omega^2_3$ on the three sphere S^3 are

$$[h_{ij}]=\begin{bmatrix}1 & 0 & 0\\ 0 & \sin^2(\theta_1) & 0\\ 0 & 0 & \sin^2(\theta_1)\sin^2(\theta_2)\end{bmatrix}.$$

Plugging these into the general formula for $\Delta_{S^{n-1}}$ above yields

$$\Delta_{S^3}\Theta=\frac{1}{\sin^2(\theta_1)}\partial_{\theta_1}\left(\sin^2(\theta_1)\partial_{\theta_1}\Theta\right)+\frac{1}{\sin^2(\theta_1)\sin(\theta_2)}\partial_{\theta_2}\left(\sin(\theta_2)\partial_{\theta_2}\Theta\right)+\frac{1}{\sin^2(\theta_1)\sin^2(\theta_2)}\partial_{\theta_3\theta_3}\Theta.$$

Now, to nondimensionalize the the Klein-Gordon equation, the substitution

$$s=\frac{ct}{f}\qquad\Psi(t,\theta)=\psi(s(t),\theta)$$

is used, and as before, a separable solution is sought. In particular, letting

$$\psi(s,\theta_1,...,\theta_{n-1})=S(s)\Theta(\theta_1,...,\theta_{n-1}),$$

the Klein-Gordon equation can be separated into the following two equations,

$$S'' - \frac{(2-n)4s^2 + (n-1)}{s(4s^2-1)}S' + \frac{(l(l+n-2) + \lambda^2 s^2)(4s^2-1)}{s^2}S = 0$$

$$\Delta_{S^{n-1}}\Theta + l(l+n-2)\Theta = 0,$$

where λ is the dimensionless parameter

$$\lambda = \frac{f\mu c}{\hbar},$$

and l is an arbitrary natural number. The term $-l(l+n-2)$ comes from the fact that in the separation of variables, the angular equation takes the form

$$\Delta_{S^{n-1}}\Theta = \beta\Theta$$

for an as yet arbitrary constant β. However the only possible values β can take - i.e. the only eigenvalues of this operator - are values of the form $-l(l+n-2)$ for l in $\mathbb{N}$.

Just as in the case of $n = 2$, the investigation into the behavior of the Klein-Gordon field across the transition boundary boils down to investigating the radial/temporal equation

$$S'' - \frac{(2-n)4s^2 + (n-1)}{s(4s^2-1)}S' + \frac{(l(l+n-2) + \lambda^2 s^2)(4s^2-1)}{s^2}S = 0 \tag{7.18}$$

Lamentably this equation can only be solved analytically for the case of $n = 2$ and $\lambda = 0$. This means that the methods used in previous chapters to investigate the behavior at the transition, which relied on having exact solutions, will not be sufficient for this more general case. To move forward in the analysis I turn to asymptotic series and global perturbation theory.

Chapter Eight: Interlude: Asymptotic Series

This chapter presents a quick introduction to asymptotic series and global perturbation theory. It presents the semi-classical approximation used in Quantum Mechanics as an example, and discusses the method of dominant balance as an approach for obtaining asymptotic approximations.

8.1 Global Perturbation Theory and Asymptotic Series

The Frobenius method is an example of a **local perturbation theory**, where an approximation to the solutions is sought, which is valid in a neighborhood of a fixed point s_0. When s_0 is an ordinary point, or a regular singular point, the approximation is in the form of a convergent series. If the point s_0 is an irregular singular point, a more general series, known as an asymptotic series is needed. Such series, which will be discussed below, often do not converge. However they can be extremely powerful, not only as approximation tools but also analytical tools.

The counterpart to local perturbation theory is **global perturbation theory**. In this realm, an approximate series solution is sought, not about any point s_0 in the domain, but rather in terms of some other small parameter in problem. The well-known semiclassical approximation $\hbar \to 0$, in quantum mechanics, is an example of global perturbation theory. Another example is the geometric optics approximation which has been considered in these investigations. Sometimes the series solutions can take the form of a Taylor series, or Frobenius series, in terms of the small parameter in question. Often, they do not. In both of the previously mentioned examples, the series approximations are asymptotic series, as in the perturbation about an irregular singular point.

For example, consider the differential equation

$$\epsilon \frac{d^2y}{ds^2} + 2\frac{dy}{ds} + y = 0,$$

where ϵ is a small parameter. The solutions to this equation can be obtained analytically. Indeed they are

$$y_1(s) = e^{\left(\frac{-1+\sqrt{1-\epsilon}}{\epsilon}\right)s}, \qquad y_2(s) = e^{\left(\frac{-1-\sqrt{1-\epsilon}}{\epsilon}\right)s},$$

and linear combinations thereof. Expanding the first of these solutions about $\epsilon = 0$ results in

$$y_1(s) = e^{-s/2} - \frac{1}{8}se^{-s/2}\epsilon + \frac{1}{128}s(s-8)e^{-s/2}\epsilon^2 + O(\epsilon^3),$$

which has the form of a Taylor series. It is not hard to believe that this series converges for any fixed s. On the other hand, the expansion of $y_2(s)$ about $\epsilon = 0$ can be written as

$$y_2(s) = e^{-\frac{2s}{\epsilon}}\left(e^{s/2} + \frac{1}{8}se^{s/2}\epsilon + \frac{1}{128}s(s+8)e^{s/2}\epsilon^2 + O(\epsilon^3)\right).$$

This series is definitely not of the Frobenius form. The meaning of such a series, and its usefulness, will be discussed shortly. However before delving into that, notice that usually (though not always), the purpose of these perturbation techniques is to find approximate solutions to problems that are too difficult to solve exactly. So having the analytic solutions available for the differential equation at hand is rarely the case. The goal is to build the approximate solutions without the having to solve the equation first. Often, a good starting place is to consider the unperturbed problem; that is the problem with the small parameter set to zero, which is typically much easier to solve. In this case the unperturbed equation is

$$2\frac{dy}{ds} + y = 0,$$

whose general solution takes the form

$$y(s) = \alpha e^{-\frac{s}{2}}.$$

This solution is the lowest order of the solution $\alpha y_1(s)$ of the perturbed equation. But all trace of the second solution $y_2(s)$ seems to have been lost. This behavior is typical of singular, global perturbation analysis. The question is, how to recover perturbation expansions to both solutions of the original differential equation?

8.2 Asymptotic Approximations and Asymptotic Series

The following introduction to asymptotic analysis will be all too brief. For a more in depth, but very accessible presentation, see Holmes [2012]. The definitions and terminology used here are borrowed from Holmes' treatment. It is particularly useful to begin the discussion with the introduction of some notation, called Landau's big-O and little-o symbols.

Given two real-valued functions $f(s)$ and $\phi(s)$, the notation

$$f = O(\phi) \quad \text{as } s \downarrow s_0$$

means that there exists constants k and s_1, independent of s, such that

$$|f(s)| \leq k|\phi(s)| \quad \text{for } s_0 < s < s_1.$$

This is Landau's big-O symbol. Similarly, given the two functions f and ϕ, the notation

$$f = o(\phi) \quad \text{as } s \downarrow s_0$$

means that *for every* $\delta > 0$ there exits and s_1, independent of s, such that

$$|f(s)| \leq \delta|\phi(s)| \quad \text{for } s_0 < s < s_1.$$

This is Landau's little-o symbol. An oft-used alternative notation for the little-o symbol is $f << \phi$ as $s \downarrow s_0$. What the big-O symbol and little-o symbols are capturing are relationships between functions that hold in a particular limit. These are very useful concepts when constructing simple approximations to complicated or unknown functions. These definitions generalize easily to the cases of $s \uparrow s_0$, and $s \to \pm\infty$, but for the sake of concreteness the limit $s \downarrow s_0$ will continue to be used in the following definitions. Also, these definitions can easily be extended to complex valued functions, or other normed linear spaces by using the appropriate norm.

Using the Landau notation, the concept of an **asymptotic approximation** can be defined as follows: given functions $f(s)$ and $\phi(s)$, we say that ϕ is an *asymptotic approximation* to f in the limit $s \downarrow s_0$, denoted

$$f \sim \phi \quad (s \downarrow s_0),$$

whenever

$$f - \phi = o(\phi) \quad \text{as } s \downarrow s_0.$$

In other words, the difference between f and ϕ, in the limit that s approaches s_0, is much smaller (in magnitude) than ϕ itself. If $\phi(s)$ is not zero near s_0, then an equivalent characterization for $f \sim \phi \quad (s \downarrow s_0)$ is

$$\lim_{s \to s_0^+} \frac{f(s)}{\phi(s)} = 1.$$

Now, a sequence of functions $\phi_1(s), \phi_2(s), ...$ is an **asymptotic sequence** as $s \downarrow s_0$ if

$$\phi_{m+1} = o(\phi_m) \quad \text{as } s \downarrow s_0$$

holds for all m. That is, every function in the sequence is much smaller in magnitude than the previous function in the sequence, in the limit $s \downarrow s_0$. The functions $\phi_1(s), \phi_2(s), ..$ are also said to be **well-ordered.**

Given such an asymptotic sequence, a function $f(s)$ is said to have an **asymptotic series expansion** to n terms if there exists n constants $\{a_1, ..., a_n\}$, independent of s, such that following asymptotic approximations hold (as $s \downarrow s_0$):

$$\begin{aligned} f - a_1\phi_1 &= o(\phi_1) \\ f - a_1\phi_1 - a_2\phi_2 &= o(\phi_2) \\ &\vdots \\ f - a_1\phi_1 - a_2\phi_2 - ... - a_n\phi_n &= o(\phi_n). \end{aligned}$$

Alternatively, this set of asymptotic relations can be written as

$$f = \sum_{k=1}^{m} a_k\phi_k + o(\phi_m) \quad \text{as } s \downarrow s_0 \quad \forall\ m \in \{1, ..n\}.$$

When this is the case, we write

$$f \sim a_1\phi_1 + a_2\phi_s + ... + a_n\phi_n \quad (s \downarrow s_0).$$

When there exists an asymptotic series expansion to n terms, for all $n \in \mathbb{N}$, then it is simply called an **asymptotic series expansion**, and is denoted as

$$f \sim \sum_{k=1}^{\infty} a_k\phi_k \quad (s \downarrow s_0).$$

The asymptotic sequence of functions used for the asymptotic series expansion are called the **scale** functions, or **basis** functions. If f has an asymptotic series expansion in terms of the scale functions $\{\phi_1(s), \phi_s(s), ...\}$, then it is unique. The first constant a_1 can be obtained by taking the limit

$$\lim_{s \downarrow s_0} \frac{f(s)}{\phi_1(s)}.$$

This limit is guaranteed to exist since by definition of the asymptotic series expansion $f - a_1\phi_1 = o(\phi_1)$. This means that for any $\delta > 0$ there exists as s_1 such that

$$\frac{|f(s) - a_1\phi_1(s)|}{|\phi_1(s)|} < \delta \quad \text{for } s_0 < s < s_1.$$

This can alternatively be written as

$$\left| \frac{f(s)}{\phi_1(s)} - a_1 \right| < \delta \quad \text{for } s_0 < s < s_1,$$

which means that the limit of $f(s)/\phi_1(s)$ as $s \downarrow s_0$ exists and is equal to a_1. Similarly a_2 can be uniquely obtained by taking the limit

$$\lim_{s \downarrow s_0} \frac{f(s) - a_1\phi_1(s)}{\phi_2(s)},$$

and so on so forth.

Asymptotic expansions, while typically not convergent can be extremely useful. Often times they provide great numerical approximations with only a few terms (see Bender and Orszag [1999] ch. 7 for interesting examples). Moreover, since asymptotic series are not restricted to strictly polynomial basis functions, they can often provide a very useful analytic approximation to the function being approximated. They are typically expressed in terms of functions that are not so complicated as the exact solution, which must contain all the detailed behavior, but not so simple as just power of x. This allows the researcher to use them in analytic explorations, so long as she is careful to manage the error terms appropriately.

8.3 Global Perturbation Series

All of the definitions above, about perturbation series and asymptotic expansions, work equally well when dealing with more than one independent variable. Simply

pick one variable in terms of which the expansion is to take place, and hold all the other variables fixed. The behavior of the series however - whether it converges or not, or how quickly it does - will in general depend on the fixed value of the other variables. This situation remains the same if instead of more than one dependent variable, the mathematical problem has a single dependent variable and one or more parameters. While the parameters are typically fixed a priori (say by the physics of the situation), mathematical expansions about a parameter are used to explore different regimes in which the behavior of solutions can be qualitatively different. Often this takes the form of comparing the parameter value to some relevant quantity in the solution, like say energy as compared to Plank's constant.

Following this train of thought, consider the time independent Schrödinger equation in one dimension

$$E\psi = -\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + V(x)\psi,$$

where V is a fixed potential function and E is the fixed energy of the state. While $\hbar$ is a fixed physical constant, the semiclassical approximation in quantum mechanics is often phrased as the limit on which $\hbar$ goes to zero. One way to make sense of this limit, is to rewrite the equation as in terms of a non-dimensionalized variable $y = x/L$ where L is a characteristic length of the system often determined by the potential $V(x)$. The equation then becomes

$$\tilde{\epsilon}\frac{d^2\psi}{dy^2} = \left(\frac{V(y)}{E} - 1\right)\psi \qquad \text{where } \tilde{\epsilon} = \frac{\hbar^2}{2mEL^2}.$$

Now, the semiclassical limit can be understood as the limit in which $\hbar/\sqrt{2mE} << L$. Note that classically, $\sqrt{2mE}$ is equal to momentum, and $\hbar$ divided by momentum is the deBroglie wavelength of the particle (up to a factor of 2π). Therefore, the semi-classical approximation is taken as the limit where the deBroglie wavelength of the particle is much smaller than the characteristic length of the problem.

Letting $V(x)/E - 1$ be denoted by $Q(x)$, it can be shown that for $Q(x) \neq 0$, the solutions to the time independent Schrödinger equation have asymptotic approximations of the form

$$\psi(x) \sim \exp\left(\pm\frac{1}{\sqrt{\tilde{\epsilon}}}\int^x \sqrt{Q(x')}\; dx'\right) Q(x)^{-1/4} \qquad (\epsilon \to 0)$$

or linear combinations thereof. These are global asymptotic approximations in the parameter ϵ. They already provide key insight into the behavior of solutions. In

particular, notice that for $Q(x) > 0$ (which means $V(x) > E$) the dominant behavior of the solutions is exponentially growing or decaying, while for $Q(x) < 0$ (which means $V(x) < E$) the dominant behavior is oscillatory. This is strongly reminiscent of the situation of a Klein-Gordon in a transition universe.

One of the main tools in building such asymptotic approximations to solutions to differential equations, both local and global, is called the **method of dominant balance**. At its core lies the idea that in a given limit, not all the terms in the differential equation contribute equally. In fact it is usually two terms that are dominant (though not always), and these two terms must therefore cancel each other (asymptotically) for the equation to be satisfied; hence the name *dominant balance*. The procedure involves

1. making an educated guess about which terms are dominant,
2. solving the associated differential equation, which is obtained by ignoring all but the dominant terms,
3. and then checking that the solution is consistent with the assumption of which terms dominate.

If it is not consistent, the choice of dominant balance was not correct, so try a different dominant balance. If it is consistent, then the solution to the associated differential equation which was just obtained is an asymptotic approximation to the solution to the original equation, in the appropriate limit. Now, if desired, the asymptotic approximation can be extended another term, by plugging into the original equation an ansatz solution consisting of the asymptotic approximation just obtained plus an undetermined function. This function is assumed to be much smaller (little-o) than the previous term. The dominant balance procedure is then repeated for this undetermined function. Doing so leads to the construction of an asymptotic series, order by order.

It might happen that the series constructed through the dominant balance method is a Taylor or Frobenius series. As previously stated these are also asymptotic series. On the other hand, it might be that the solution has an essential singularity at the limit being considered. These solutions often behave like the exponential of a singular function near the singularity. It is for this reason that usually the first step

in constructing an asymptotic series solution, around singular limits, is to perform the substitution

$$y(s) = e^{S(s)}.$$

Generally, this makes the dominant balance procedure simpler.

For example, consider the equation

$$\epsilon \frac{d^2y}{ds^2} + 2\frac{dy}{ds} + y = 0,$$

which was analyzed at the beginning of this section. The limit ϵ goes to zero is indeed a singular limit. A tell-tale sign of this behavior is the fact that the equation changes order when ϵ is set exactly to zero. This suggests that performing the transformation $y = e^S$ is a good idea, which leads to the following differential equation for S:

$$\epsilon \left(S'' + (S')^2\right) + 2S' + 1 = 0.$$

There are $\binom{4}{2}$ possible choices for two-term dominant balances. One such choice is that the first and last terms dominate, that is

$$\epsilon S'' \sim -1 \qquad (\epsilon \downarrow 0),$$

and the other terms are $o(1)$ as $\epsilon \downarrow 0$. Solving the associated equation, which is obtained by replacing $\sim$ with $=$, leads to

$$S = -\frac{1}{2\epsilon}s^2 + as + b.$$

Note however that this would mean that the term $\epsilon(S')^2$ would be asymptotic to ϵ^{-1}. This is in contradiction with the choice of dominant balance, whereby the assumption was that $\epsilon(S')^2$ is $o(1)$. So this is an inconsistent dominant balance.

Another possible dominant balance is

$$\epsilon(S')^2 \sim -2S'.$$

The solution to the associated differential equation is

$$S(s) = -\frac{2s}{\epsilon} + a.$$

This means that the two terms that were kept are asymptotic to ϵ^{-1}, while the two terms that were dropped are asymptotic to 1. Therefore this is a consistent dominant balance. To find the next term in the asymptotic series for S, the solution

$$S(s) = -\frac{2s}{\epsilon} + A(s, \epsilon)$$

where

$$A = o\left(\epsilon^{-1}\right) \quad (\epsilon \downarrow 0)$$

is plugged into the original equation. This leads to

$$\epsilon\left(A'' + (A')^2\right) - 2A' + 1 = 0.$$

The only consistent dominant balance turns out to be

$$2A' \sim 1,$$

which leads to $A = s/2 + B(\epsilon, s)$, where $B = o(1)$. So the solution to the original equation has an asymptotic expansion

$$y(s) = e^{-\frac{2s}{\epsilon}} e^{s/2}\left(1 + o(1)\right) \qquad (\text{as } \epsilon \downarrow 0),$$

which is exactly the first term of the series expansion of the exact solution

$$y_2(s) = e^{\left(\frac{-1-\sqrt{1-\epsilon}}{\epsilon}\right)s}.$$

In this particular situation, the approximation to the other solution $y_1(s)$, would come from a different choice of initial dominant balance, namely

$$2S' \sim -1.$$

Chapter Nine: Klein-Gordon - Approach 3

This chapters presents a solution to the problem of matching local solutions of the Klein-Gordon equation of general mass in a transition spacetime of general dimensions. First, local approximations about the points $s = 0$ and $s = 1/2$ are obtained and then a global approximation, in the geometric optics limit, is used to bridge the two local approximations. Matching local to global will require the use of an intermediate approximation and a non-standard method of matching detailed in appendix C. The relevant regimes studied in this chapter are highlighted in the roadmap picture 9.1.

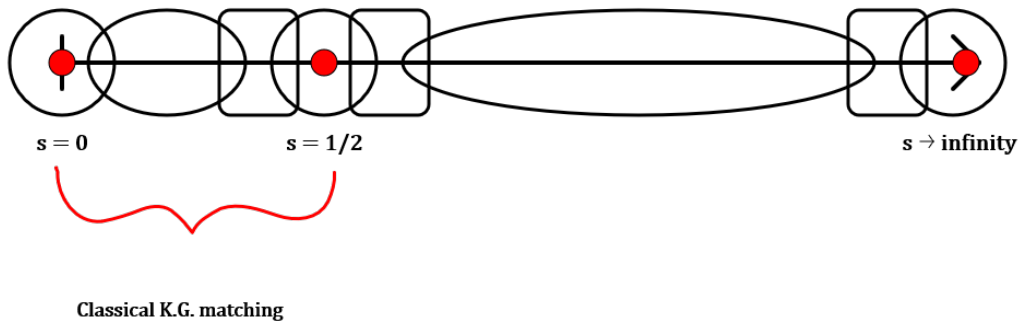


Figure 9.1: These are the regimes involved in the classical matching problem.

9.1 The General Idea

With the tools of asymptotic expansions, global perturbation analysis, and the method of dominant balance, it is now possible to answer the question of whether or not reality and regularity are mutually exclusive conditions for the general n and λ transition-universe Klein-Gordon. The strategy is to obtain local approximations of solutions

for $s = 0$ and $s = 1/2$, and match them via a global perturbation approximation of solutions. This way, imposing regularity at $s = 0$ will translate into a particular combination of the global (approximate) solutions, which can then be matched to a particular linear combination of local (approximate) solutions at $s = 1/2$. The global approximation sought will be in the limit of geometric optics described in section 3.1. In the context of quantum mechanics and the semiclassical approximation, this method is known as the WKB approximation, where the global $\hbar \to 0$ approximation is matched to local approximations at the classical turning points.

The basic idea behind this procedure is that the differential equation for S has two linearly independent solutions over the entire domain, even if the equation cannot be solved analytically. For concreteness, denote (a choice of) these solutions as $S_1(s)$ and $S_2(s)$. Furthermore, any other solution to the differential equation must be expressible as a linear combination of $S_1(s)$ and $S_2(s)$. The local approximations so far obtained, say about $s = 0$, are local expansions for solutions of the differential equation. Since two linearly independent expansions were obtained, they are local expansions to two linearly independent solutions, albeit not necessarily of $S_1(s)$ and $S_2(s)$. Instead, they will generally be local expansions of two different (independent) combinations of $S_1(s)$ and $S_2(s)$, say

$$S_3(s) = a_1 S_1(s) + a_2 S_2(s), \quad S_4(s) = b_1 S_1(s) + b_2 S_2(s).$$

In chapter 6, where the exact solutions $S_1(s)$ and $S_2(s)$ could be obtained, determining the constants a_1, b_2, b_1, b_2 of the combinations above, was relatively straightforward. First the analytic solutions $S_1(s)$ and $S_2(s)$ were expanded about the point in question, in this case $s = 0$. Then the linear combination of these expansions that matched the expansions obtained directly from the Frobenius method became evident.

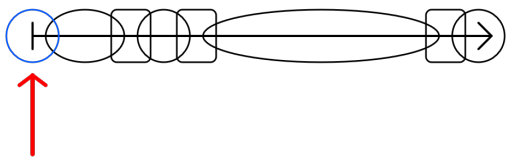
Once the constants a_1, a_2, b_1, and b_2 were determined, relating the analytic solutions, to the local solutions about $s = 0$, the same procedure could be done to relate the analytic solutions, to local solutions about $s = 1/2$. This determined another set of constants, say c_1, c_2, d_1, d_2, specifying these particular linear combinations. This allowed the constants a_1, a_2, b_1, b_2 to be expressed in terms of the constants c_1, c_2, d_1, d_2, or vice-versa. So, in particular, the effects of imposing a boundary condition at $s = 0$ on the solutions at $s = 1/2$ (or vice-versa) could be understood.

In this chapter, on the other hand, the analytic solutions $S_1(s)$ and $S_2(s)$ are not known. These were crucial in connecting local solutions because they are not

local themselves. However, global approximations, like the ones computed in section 8.3, can equally serve as this bridge. The global approximations themselves are approximating linear combinations of the solutions $S_1(s)$ and $S_2(s)$. Indeed they are global approximations to two linearly independent solutions, which could equally well be used to relate local solutions at $s = 0$ with local solutions at $s = 1/2$. The key, however, is that these exact solutions are not needed; their global approximations suffice.

9.2 Local Approximations

Local $s \to 0$ Approximation



The radial/temporal part of the Klein-Gordon equation (with separable solution ansatz) for the toy transition universe of dimension n and generic effective mass λ is

$$S'' - \frac{(2-n)4s^2 + (n-1)}{s(4s^2-1)}S' + \frac{(l(l+n-2) + \lambda^2 s^2)(4s^2-1)}{s^2}S = 0. \tag{9.1}$$

As previously stated, the equation for S above cannot be solved analytically for any choice of n and λ, other than $n = 2$ and $\lambda = 0$. Therefore to analyze the behavior of the solutions as s approaches zero, I turn to local perturbation analysis as done in chapter 6. Equation (9.1) can be rewritten as

$$S'' + \frac{p(s)}{s}S' + \frac{q(s)}{s^2}S = 0,$$

where

$$p(s) = \frac{(n-2)4s^2 + (1-n)}{(4s^2-1)}, \qquad q(s) = (l(l+n-2) + \lambda^2 s^2)(4s^2-1).$$

Since p and q are both regular at $s = 0$, the point $s = 0$ is a *regular singular point*.

The indicial equation becomes

$$\alpha^2 + (n-2)\alpha - l(l+n-2) = 0,$$

which can be rewritten as

$$(\alpha - l)(\alpha + l + n - 2) = 0.$$

So the solutions are $\alpha_1 = l$ and $\alpha_2 = -(l+n-2)$. Notice the perhaps not so surprising fact that the pair of solutions are exactly the product decomposition of the eigenvalues to the Laplace-Beltrami operator $\Delta_{S^{n-1}}$.

Now, the difference in solutions to the indicial equation is a nonzero integer:

$$\Delta\alpha = (2l + n - 2),$$

since l must be an integer greater than zero. This means that only one of the two linearly independent solutions defined in a neighborhood of $s = 0$ is guaranteed to have a Frobenius series expansion. For $n > 2$, the term $-(l+n-2)$ is always negative, so the the indicial exponent l is greater than the indicial exponent $-(l + n - 2)$. Therefore the solution with a guaranteed Frobenius series expansion can be expressed as

$$y_l(s) = s^l \sum_{k=0}^{\infty} a_k s^k.$$

This solution is regular as s approaches zero. Indeed the first few terms of the expansion are

$$y_l(s) = s^l \left(1 - \frac{4l(l+n-1) + \lambda^2}{2(2l+n)} s^2 + O(4).\right)$$

The other linearly independent solution might be a little bit more complicated. The answer depends on whether or not the RHS of the recursion relation

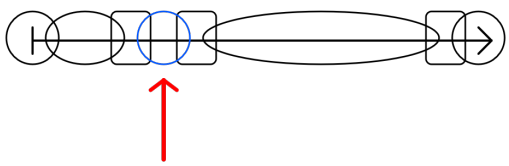
$$\left((\alpha_2 + k)^2 + (p_0 - 1)(\alpha_2 + k) + q_0\right) a_k = -\sum_{j=0}^{k-1} \left((\alpha_2 + j)p_{k-j} + q_{k-j}\right) a_j \qquad \text{for } k = 1, 2, ...$$

also vanishes when the LHS vanishes (when $k = 2(l+1)$). However, for the purposes of this investigation, it is not necessary to know the explicit expansion of the solution associated with the indicial exponent $-(l + n - 2)$. It is sufficient to know that this solution will blow up at least as fast as

$$s^{-(l+n-2)}$$

as s approaches zero (see Bender and Orszag [1999]).

Local $s \to 1/2$ Approximation



To obtain the local approximations as s approaches $1/2$, equation (9.1) is rewritten as

$$S'' + \frac{p(s)}{\left(s-\frac{1}{2}\right)}S' + \frac{q(s)}{\left(s-\frac{1}{2}\right)^2}S = 0,$$

where

$$p(s) = \frac{(n-2)4s^2 + (1-n)}{4s\left(s+\frac{1}{2}\right)}, \qquad q(s) = \frac{4(l(l+n-2)+\lambda^2 s^2)\left(s+\frac{1}{2}\right)\left(s-\frac{1}{2}\right)^3}{s^2}.$$

Again, since p and q are regular about $s = 1/2$, this is a *regular singular point*. For these coefficients the indicial equation becomes

$$\alpha^2 - \frac{3}{2}\alpha = 0,$$

which, surprisingly, is exactly the same one as for the case of $n = 2$ and $\lambda = 0$. This means that at least to lowest order, the solutions will behave exactly the same as they did for the case of $n = 2$ and $\lambda = 0$. In particular, the two linearly independent solutions $y_{3/2}(s)$ and $y_0(s)$ will have local expansions of the form

$$y_{3/2}(s) = \left(s-\frac{1}{2}\right)^{3/2}\sum_{k=0}^{\infty} a_k\left(s-\frac{1}{2}\right)^k,$$

and

$$y_0(s) = \sum_{k=0}^{\infty} b_k\left(s-\frac{1}{2}\right)^k.$$

Furthermore, these Frobenius series representations are guaranteed to converge in a neighborhood of $s = 1/2$.

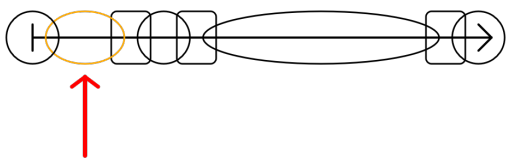
The first few terms of the expansions, as obtained through the Frobenius method described in ch 6, are

$$\begin{aligned} y_{3/2}(s) =& \left(s-\frac{1}{2}\right)^{3/2} - \left(\frac{9}{10}+\frac{6}{5}(n-2)\right)\left(s-\frac{1}{2}\right)^{5/2} + \left(\frac{69}{56}+\frac{15}{7}(n-2)+\frac{6}{7}(n-2)^2\right)\left(s-\frac{1}{2}\right)^{7/2} \\ &+ O\left(s-\frac{1}{2}\right)^{9/2} \\ y_0(s) =& 1 - \left(\frac{32}{9}l(l+n-2)+\frac{8}{9}\lambda^2\right)\left(s-\frac{1}{2}\right)^3 + O\left(s-\frac{1}{2}\right)^4, \end{aligned} \tag{9.2}$$

As before, $y_0(s)$ is real throughout the transition, and $y_{3/2}(s)$ picks up a factor if i.

So the situation for the case of general n and λ is very similar to the case of $n = 2$ and $\lambda = 0$ considered earlier. As s approaches zero, one of the solutions goes to zero and the other one blows up. Furthermore, in a neighborhood of the transition $s = 1/2$, both solutions are finite, one remaining real throughout and the other picking up a factor of i as it transitions. However, since the analysis of the solutions has only been local, either about $s = 0$ and separately about $s = 1/2$, it is difficult to see how these local solutions relate to each other. While it seems plausible (and even likely) that the regularity vs reality situation holds for this more general case, to know whether or not this is the case, a way to connect the local solutions is needed. For this I turn to global perturbation theory; in particular the WKB approximation/gometric optics limit, which is very successful in the realm of tunneling.

9.3 Global Approximation



Obtaining Global Geometric Optics Approximations

Starting from the equation governing S

$$S'' - \frac{(2-n)4s^2+(n-1)}{s(4s^2-1)}S' + \frac{(l(l+n-2)+\lambda^2 s^2)(4s^2-1)}{s^2}S = 0,$$

the first step is to factor out the $l(l+n-2)$, and for convenience, label that term ϵ^{-2}. This yields

$$\epsilon^2\left(S'' - \frac{(2-n)4s^2+(n-1)}{s(4s^2-1)}S'\right) + \frac{(1+\epsilon^2\lambda^2 s^2)(4s^2-1)}{s^2}S = 0.$$

The geometric optics limit being considered can be expressed as the limit of l being very large, or alternatively, $\epsilon \to 0$. Since the unperturbed ($\epsilon = 0$) equation is of different order, the perturbation expansion about $\epsilon = 0$ will likely not be a Taylor or Frobenius series. Therefore, the transformation

$$S(s) = e^{A(s)},$$

is probably going to make things easier. Plugging this transformation into the above equation yields

$$\epsilon^2\left(A'' + (A')^2 - \frac{(2-n)4s^2+(n-1)}{s(4s^2-1)}A'\right) + \frac{(4s^2-1)}{s^2} + \epsilon^2\lambda^2(4s^2-1) = 0.$$

After trying many different dominant balances, the consistent one turns out to be

$$\epsilon^2(A')^2 \sim -\frac{(4s^2-1)}{s^2}.$$

However, it comes with a caveat which will become clear momentarily. The solutions to the associated differential equation

$$\epsilon^2(A')^2 = -\frac{(4s^2-1)}{s^2} \qquad (\epsilon \downarrow 0)$$

are

$$A = \pm\frac{1}{\epsilon}\left(\sqrt{1-4s^2} - \tanh^{-1}\left(\sqrt{1-4s^2}\right)\right),$$

plus an arbitrary constant. To check for consistency, this solution for A is plugged into to all the dropped terms in the proposed dominant balance:

$$\epsilon^2 A'' = \epsilon\left(\frac{\mp 1}{s^2\sqrt{1-4s^2}}\right) \sim \epsilon$$

$$-\epsilon^2\frac{(2-n)4s^2+(n-1)}{s(4s^2-1)}A' = \epsilon\left(\frac{\pm((2-n)4s^2+(n-1))}{s^2\sqrt{1-4s^2}}\right) \sim \epsilon$$

$$\epsilon^2\lambda^2(4s^2-1) \sim \epsilon^2$$

where all of these asymptotic relations would hold in the limit $\epsilon \downarrow 0$. Comparing these to the terms that were kept, which are ~ 1 in the same limit, shows that indeed this is a consistent dominant balance.

The all important caveat however, is that this dominant balance is not uniformly valid over the entire s interval. As s gets closer and closer to $1/2$, two of the dropped terms will become more and more dominant. In fact, for any fixed ϵ there will be a point s close to $1/2$ at which these terms dominate, which means the asymptotic approximation derived from this dominant balance no longer holds. This does not present a big problem in this situation since we have local solutions about $s = 1/2$ (and at $s = 0$) that are very good approximations. Furthermore, it will be relatively simple to match the global asymptotic approximations to the local frobenius approximations in a region where both approximations hold.

Notice that the approximation obtained for A

$$A \sim \pm\frac{1}{\epsilon}\left(\sqrt{1-4s^2} - \tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)$$

is exactly the (phase of the) solution for the case of $n = 2$ and $\mu = 0$ obtained in section 4.1. Adding this to the fact that the local approximations about $s = 0$ and $s = 1/2$, to highest order, also behave exactly like the $n = 2$ and $\mu = 0$ local solutions, it is becoming evident that the situation is entirely analogous. In particular the question about regularity vs reality is likely to be exactly the same. However, to be entirely sure, I will continue onwards, and actually perform the matching.

To this end, it is necessary to extend the asymptotic approximation a little further. This is the case because, while

$$A \sim \pm\frac{1}{\epsilon}\left(\sqrt{1-4s^2} - \tanh^{-1}\left(\sqrt{1-4s^2}\right)\right) \qquad (\epsilon \downarrow 0),$$

it need not be the case that

$$S(s) \sim e^{\pm\frac{1}{\epsilon}\left(\sqrt{1-4s^2} - \tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)} \qquad (\epsilon \downarrow 0). \tag{9.3}$$

That depends on the sub-leading term in the asymptotic expansion to A. Indeed if the next term doesn't approach 0 as $\epsilon \to 0$, then $S(s)$ would not be asymptotic to the exponential of the one term asymptotic expansion to A. This can be checked using the definition of an asymptotic approximation. So plugging the ansatz

$$A = \pm\frac{1}{\epsilon}\left(\sqrt{1-4s^2} - \tanh^{-1}\left(\sqrt{1-4s^2}\right)\right) + B(\epsilon, s)$$

where B is much smaller than the leading term in A, in the limit $\epsilon \to 0$, results in the following equation for B:

$$\epsilon^2 \left(B'' + (B')^2 + \frac{(2-n)4s^2 + (n-1)}{s(1-4s^2)} B' \right) \pm 2\epsilon \left(\frac{\sqrt{1-4s^2}}{s} \right) B'$$
$$\pm \epsilon(n-2) \left(\frac{\sqrt{1-4s^2}}{s^2} \right) - \epsilon^2 \lambda^2 (1-4s^2) = 0.$$

The consistent dominant balance, away from $s = 0$ and $s = 1/2$ is

$$2\epsilon \left(\frac{\sqrt{1-4s^2}}{s} \right) B' \sim -\epsilon(n-2) \left(\frac{\sqrt{1-4s^2}}{s^2} \right).$$

However, this is only true for $n > 2$. For the case of $n = 2$, the RHS of the above relation vanishes, and the proper dominant balance is

$$2\epsilon \left(\frac{\sqrt{1-4s^2}}{s} \right) B' \sim \pm \epsilon^2 \lambda^2 (1-4s^2).$$

Focusing on the case of $n > 2$, the associated differential equation is

$$B' = -\frac{(n-2)}{2s},$$

whose solution is

$$B = -\frac{(n-2)}{2} \ln(s),$$

plus an arbitrary constant. So altogether, these results can be written as

$$A = \pm \frac{1}{\epsilon} \left(\sqrt{1-4s^2} - \tanh^{-1} \left(\sqrt{1-4s^2} \right) \right) - \frac{(n-2)}{2} \ln(s) + C(\epsilon, s)$$

where $C = o(1)$ as $\epsilon \downarrow 0$. Now it is possible to say that

$$S(s) \sim e^{\pm \frac{1}{\epsilon} \left(\sqrt{1-4s^2} - \tanh^{-1} \left(\sqrt{1-4s^2} \right) \right) - \frac{(n-2)}{2} \ln(s)} \qquad (\epsilon \downarrow 0),$$

since the ignored part of A is $o(1)$ as $\epsilon \downarrow 0$. With this asymptotic approximation in hand, the matching can be performed. For the sake of completion, the next two terms in the asymptotic approximation have been calculated using the method above. Altogether, the four term asymptotic approximation for $S(s)$ in the limit of $\epsilon \to 0$ is

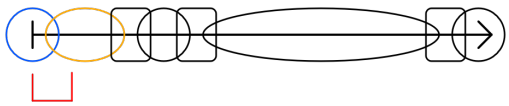
$$S(s) \sim e^{\frac{1}{\epsilon} A + B + \epsilon C + \epsilon^2 D + o(\epsilon^2)} \qquad (\epsilon \downarrow 0) \tag{9.4}$$

where

$$A = \pm\left(\sqrt{1-4s^2} - \tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)$$
$$B = -\frac{(n-2)}{2}\ln(s)$$
$$C = \pm\left(-\frac{(n-2)^2}{8}\tanh^{-1}\left(\sqrt{1-4s^2}\right) + \frac{(n-2)}{4}\frac{1}{\sqrt{1-4s^2}} - \frac{\lambda^2}{24}(1-4s^2)^{3/2}\right)$$
$$D = \left(-\frac{(n-2)}{8(4s^2-1)^2} + \frac{(n-2)(n-4)}{16(4s^2-1)} - \frac{\lambda^2}{2}(4s^2-1)\right)$$

9.4 Matching Local to Global

Matching $\epsilon \to 0$ with $s \to 0$



Through the method of dominant balance, global approximations in the small ϵ limit were obtained for two linearly independent solutions. Label these hypothetical analytic solutions $S_+(s)$ and $S_-(s)$. The two term asymptotic approximations obtained are (combined into one expression using $\pm$)

$$S_\pm(s) \sim e^{\pm\frac{1}{\epsilon}\left(\sqrt{1-4s^2}-\tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)-\frac{(n-2)}{2}\ln(s)} \qquad (\epsilon \downarrow 0) \tag{9.5}$$

Now, by expanding this asymptotic relations in the limit of $s = 0$, it can be matched to the solutions whose local behaviors were obtained using the Frobenius method in section 9.2. The expansion of the right hand side of equation (9.5), about the point $s = 0$ is

$$\exp\left(\pm\frac{1}{\epsilon}\right) s^{\pm\frac{1}{\epsilon}-\frac{(n-2)}{2}}\left(1 \mp \frac{s^2}{\epsilon} + O(s)^4\right)$$

The tell-tale terms in this expansion, that can be used to quickly match to the local expansions about $s = 0$ is the term

$$s^{\pm\frac{1}{\epsilon}-\frac{(n-2)}{2}}.$$

Recalling the definitions of ϵ as $(l(l+n-2))^{-1/2}$, this term can be rewritten as

$$s^{\pm l\sqrt{1+\frac{(n-2)}{l}}-\frac{(n-2)}{2}}.$$

Now, note that the limit $\epsilon \to 0$ is the same as the limit $l \to \infty$, therefore the radical can be expanded to obtain

$$s^{\pm\left(l+\frac{(n-2)}{2}\right)-\frac{(n-2)}{2}+O(1/l)^2}.$$

or equivalently

$$s^l \quad \text{and} \quad s^{-(l+n-2)}$$

up to order $O(l^{-1})^2$ terms in the exponent. It is expected that this approximation does not contain the exact behavior for higher orders of l^{-1}, since only terms of order ϵ^{-1} and ϵ^0 in the initial asymptotic expansion were used. Terms of $O(l^{-1})^2$ are of $O(\epsilon)^1$. Indeed, starting from the four term asymptotic approximation instead of only the two term asymptotic approximation, and expanding about $s \to 0$ as above yields the term

$$s^{\pm\frac{1}{\epsilon}-\frac{(n-2)}{2}\pm\frac{(n-2)^2}{8}\epsilon}$$

instead. Expressing this in terms of l and expanding as $l \to \infty$ yields

$$s^{\pm\left(l+\frac{(n-2)}{2}\right)-\frac{(n-2)}{2}+O(1/l)^3}.$$

That is, the $O(l^{-1})^2$ terms exactly cancel out. The higher order terms must cancel out order by order to exactly reproduce the local frobenius behavior.

The main point however, is that this hints to the fact that the solution $S_+(s)$ is a just a constant multiple of the solution with the local frobenius expansion

$$y_l(s) = s^l\left(1-\frac{4l(l+n-1)+\lambda^2}{2(2l+n)}s^2+O(s)^4\right).$$

The last step in verifying this is to make sure that the term

$$\left(1-\frac{4l(l+n-1)+\lambda^2}{2(2l+n)}s^2+O(s)^4\right)$$

is really

$$\left(1-\frac{s^2}{\epsilon}+O(s)^4\right)$$

to order ϵ^0, as obtained from the $s \to 0$ expansion of the asymptotic approximation above. Solving the relationship $l(l+n-2) = \epsilon^{-2}$ in terms of l, and expanding to $\mathrm{O}(\epsilon)^0$ yields

$$l = \pm\frac{1}{\epsilon} - \frac{(n-2)}{2} + O(\epsilon).$$

Plugging the positive solution in to the expression containing the l terms, and again keeping only terms up to $O(\epsilon)^0$ yields

$$y_l(s) = s^l \left(1 - \frac{s^2}{\epsilon} + O(s)^4 + O(\epsilon)\right)$$

as expected. So the solution S_+ is a constant multiple of the solution whose local frobenius expansion has indicial exponent l. Similarly, the solution $S_-(s)$ has

$$s^{-(l+n-2)}$$

as the leading term, which indicates the the solution $S_-(s)$ contains some of the solution whose local expansion has indicial exponent $-(l+n-2)$. Among other things, this means that this solution blows up as s approaches 0.

Matching $\epsilon \to 0$ with $s \to 1/2$

Now, to complete the overall goal, the global approximation must be matched with the local approximations for $s = 1/2$. Performing this matching turns out to be more complicated than the previous matching. Simply expanding the asymptotic approximations to $S_+(s)$ and $S_-(s)$ in the limit of $s \to 1/2$ is too crude of a technique to perform a proper matching. This can be seen by expanding the two term asymptotic approximation to $S\pm(s)$ about $s = 0$. This yields

$$\begin{aligned} e^{\pm\frac{1}{\epsilon}\left(\sqrt{1-4s^2}-\tanh^{-1}\left(\sqrt{1-4s^2}\right)\right)-\frac{(n-2)}{2}\ln(s)} &= 2^{(n-2)/2}\left(1-(n-2)\,(s-1/2) \pm \frac{i8}{3\epsilon}\left(s-\frac{1}{2}\right)^{3/2}\right. \\ &+ \left((n-2)+\frac{(n-2)^2}{2}\right)\left(s-\frac{1}{2}\right)^2 \mp \frac{4i}{15\epsilon}(9+10(n-2))\left(s-\frac{1}{2}\right)^{5/2} \\ &\left.- \left(\frac{32}{9\epsilon^2}+\frac{4(n-2)}{3}+(n-2)^2+\frac{(n-2)^3}{6}\right)\left(s-\frac{1}{2}\right)^3 + O\left(s-\frac{1}{2}\right)^{7/2}\right). \end{aligned}$$

Recalling the the local (approximation to) solutions about $s = 1/2$ using the Frobenius method

$$\begin{aligned} y_{3/2}(s) =& \left(s-\frac{1}{2}\right)^{3/2} - \left(\frac{9}{10}+\frac{6}{5}(n-2)\right)\left(s-\frac{1}{2}\right)^{5/2} + \left(\frac{69}{56}+\frac{15}{7}(n-2)+\frac{6}{7}(n-2)^2\right) \\ & \times \left(s-\frac{1}{2}\right)^{7/2} + O\left(s-\frac{1}{2}\right)^{11/2} \\ y_0(s) =& 1 - \left(\frac{32}{9}l(l+n-2)+\frac{8}{9}\lambda^2\right)\left(s-\frac{1}{2}\right)^3 + O\left(s-\frac{1}{2}\right)^4, \end{aligned}$$

it is clear that the $s \to 1/2$ expansion of the two terms asymptotic approximation cannot be a linear combination of $y_{3/2}$ and y_0. In particular the global and then local expansion has a term proportional to $(s-1/2)$ that is not in either local series approximation. However, keeping only the most dominant terms in the $\epsilon \to 0$ limit of the global and then local expansion yields

$$\begin{aligned} & \pm 2^{(n-2)/2}\frac{i8}{3\epsilon}\left(\left(s-\frac{1}{2}\right)^{3/2} - \left(\frac{9}{10}+(n-2)\right)\left(s-\frac{1}{2}\right)^{5/2}\right) - \frac{2^{(n-2)/2}}{\epsilon^2}\left(\frac{32}{9}\right)\left(s-\frac{1}{2}\right)^3 \\ & + O\left(s-\frac{1}{2}\right)^{7/2}. \end{aligned} \tag{9.6}$$

This suggests that it is not entirely far off from the correct expansion, and furthermore, that both S_+ and S_- are likely to be linear combinations of the solutions whose local frobenius expansions have been obtained.

The lack of matching between the global $\epsilon \to 0$ approximation (then expanded locally about $s = 1/2$) and the local $s = 1/2$ frobenius series solutions arises from the fact that the global approximation is not uniformly valid over the entire s domain. Indeed, when the transformation $S(s) = e^{A(s)}$ was plugged into the exact equation for $S(s)$, to obtain

$$\epsilon^2\left(A'' + (A')^2 - \frac{(2-n)4s^2+(n-1)}{s(4s^2-1)}A'\right) + \frac{(4s^2-1)}{s^2} + \epsilon^2\lambda^2(4s^2-1) = 0,$$

the proposed dominant balance was

$$\epsilon^2(A')^2 \sim -\frac{(4s^2-1)}{s^2}.$$

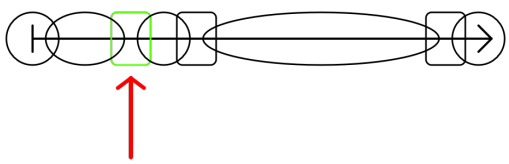
In particular, this assumes that both of these terms are much larger than

$$-\epsilon^2\frac{(2-n)4s^2+(n-1)}{s(4s^2-1)}A'.$$

While this is true for small ϵ and moderate s, it is not true for s arbitrarily close to $1/2$. For any fixed small but fixed ϵ, there is a neighborhood of s in which this term will generally be as large or larger than the terms in the proposed dominant balance, thus making the dominant balance inconsistent.

The solution to this dilemma is to obtain an *intermediate approximation*, keeping terms that matter when both limits, $\epsilon \to 0$ and $s \to 1/2$ are taken simultaneously. This necessarily leads to a more complicated approximation that is hopefully valid in a domain that intersects with each of the simpler approximations that are being matched. If such an approximation can be found, it would then be used to bridge between the two approximation whose matching is sought. In the case at hand, it would be matched on one side to the $\epsilon \to 0$ approximation and on the other side to the $s \to 1/2$ approximation. This is effectively playing the same game that is already being played by matching the $s \to 0$ and $s \to 1/2$ approximations via the global $\epsilon \to 0$ approximation.

Matching Via Intermediate Approximation



To obtain this intermediate approximation, both limits $\epsilon \to 0$ and $s \to 1/2$ must be considered simultaneously. The first step is to determine the relative rate at which the limits should be taken. This is often achieved by performing what is called a **stretching transformation**, which initially leaves this relative rate undetermined, and then finding a **distinguished limit** of the equation. This is the limit in which the desired terms are all relevant.

In particular, for the equation at hand, the stretching transformation amounts to the change of coordinates from s to

$$r = \frac{1}{\epsilon^\alpha}\left(\frac{1}{2} - s\right)$$

where the constant α is greater than zero, but is otherwise undetermined. The $\epsilon \to 0$ limit will be taken, while holding r fixed. This implies that s will tend to $1/2$ also.

Therefore, in the r coordinates, the global $\epsilon \to 0$ approximation is equivalent to the mixed global $\epsilon \to 0$ and local $s \to 1/2$ approximation in the s coordinate. In effect, this transformation maps a small neighborhood of $s = 1/2$ to the entire r domain, which is where it gets the name *stretching transformation*.

Now, the relative rate at which s goes to $1/2$ as compared to how fast ϵ goes to zero is controlled by the constant α. This is chosen so that the dominant balance includes the most relevant terms when the $\epsilon \to 0$ and $s \to 1/2$ limits are considered independently. In particular, starting with the equation for $S(s)$ (slightly rewritten)

$$S'' + \left(\frac{(n-2)}{s} + \frac{1}{s(1-4s^2)} \right) S' - \left(\frac{1}{\epsilon^2} \frac{(1-4s^2)}{s^2} + \lambda^2 (1-4s^2) \right) S = 0$$

the dominant balance in the $\epsilon \to 0$ limit consists of the first and fourth terms, that is

$$S'' \sim \frac{1}{\epsilon^2} \frac{(1-4s^2)}{s^2} S.$$

On the other hand the dominant balance in the $s \to 1/2$ limit consists of the first and third terms

$$S'' \sim -\frac{1}{s(1-4s^2)} S'.$$

Therefore, the relevant mixed local and global limit is one in which all three terms

$$S'', \quad \frac{1}{\epsilon^2} \frac{(1-4s^2)}{s^2} S, \quad -\frac{1}{s(1-4s^2)} S' \tag{9.7}$$

balance out.

After the change of coordinates to r is performed, these three terms will correspond to three terms in the new equation. So, α will be chosen in such a way that all three of these terms contribute. This is the *distinguished limit* referred to earlier.

Performing the transformation

$$s = \frac{1}{2} - \epsilon^\alpha r, \qquad S(s) = U(r(s))$$

leads to

$$U'' + P(r,\epsilon) U' + Q(r,\epsilon) U = 0$$

where

$$\begin{aligned} P(r,\epsilon) &= -\left(\frac{2(n-2)\epsilon^\alpha}{1-2r\epsilon^\alpha} + \frac{1}{2r(1-r\epsilon^\alpha)(1-2r\epsilon^\alpha)} \right) \\ Q(r,\epsilon) &= -\left(\frac{16 r \epsilon^{3\alpha-2}(1-r\epsilon^\alpha)}{(1-2r\epsilon^\alpha)^2} + 4\lambda^2 r \epsilon^{3\alpha} - 4\lambda^2 r^2 \epsilon^{4\alpha} \right). \end{aligned}$$

The coefficients $P(r,\epsilon)$ and $Q(r,\epsilon)$ have the following expansions in ϵ, as $\epsilon \to 0$

$$\begin{aligned} P(r,\epsilon) =& -\frac{1}{2r} - \left(\frac{3}{2} + 2(n-2)\right)\epsilon^{\alpha} + O(\epsilon)^{4\alpha} \\ Q(r,\epsilon) =& -16r\epsilon^{3\alpha-2} - 48r^2\epsilon^{4\alpha-2} + O(\epsilon)^{5\alpha-2} + O(\epsilon)^{3\alpha} \end{aligned}$$

where the second expansion contains two big-O terms since the well-ordering of terms is not yet known. It is clear that the most dominant term is of order $\epsilon^{3\alpha-2}$, but without a specification of α it is not clear whether $\epsilon^{4\alpha-2}$ dominates $\epsilon^{3\alpha}$, or the other way around. Similarly for the term of order $\epsilon^{5\alpha-2}$ and so on so forth.

Plugging (the most dominant terms of) these expansions into the equation for U leads to

$$U'' - \left(\frac{1}{2r} + O(\epsilon)^{\alpha}\right)U' - \left(16r\epsilon^{3\alpha-2} + O(\epsilon)^{4\alpha-2} + O(\epsilon)^{3\alpha}\right)U = 0.$$

The three terms that must balance out in the limit $\epsilon \to 0$ are

$$U'', -\frac{1}{2r}U', -16r\epsilon^{3\alpha-2}U.$$

These three terms correspond to the three terms in equation (9.7) via the stretching transformation (at least their most dominant parts). So $3\alpha - 2$ must equal to zero. The distinguished limit is therefore $\alpha = 2/3$. In this case the dominant balance is

$$U'' - \frac{1}{2r}U' \sim 16rU.$$

Ultimately this is the choice that balances the global $\epsilon \to 0$ limit with the $s \to 1/2$ limit.

The associated differential equation has two linearly independent solutions

$$e^{\pm\frac{8}{3}r^{3/2}}.$$

These represent the dominant behavior of two linearly independent exact solutions, in the mixed global and local limit determined by $\alpha = 2/3$. Having chosen a *distinguished limit* $\alpha = 2/3$, obtaining the next term in the asymptotic approximation is just an application of the dominant balance method in the limit $\epsilon \to 0$, while being careful to include enough terms of the expansions of the coefficients P and Q.

Plugging in the ansatz

$$U = \exp\left(\pm\frac{8}{3}r^{3/2} + B\right)$$

into the equation for U, keeping terms of up to order $\epsilon^{2/3}$, leads to

$$B''+(B')^2+\left(\pm 8r^{1/2}-\frac{1}{2r}+\left(\frac{5}{2}-2n\right)\epsilon^{2/3}\right)B'+\left(-48r^2\pm 4\left(\frac{5}{2}-2n\right)r^{1/2}\right)\epsilon^{2/3}=0.$$

It turns out that there is no consistent two-term dominant balance, therefore a three term dominant balance must be considered, namely

$$B''+\left(\pm 8r^{1/2}-\frac{1}{2r}\right)B'+\left(-48r^2\pm 4\left(\frac{5}{2}-2n\right)r^{1/2}\right)\epsilon^{2/3}=0.$$

Now, the equation above can be solved analytically, however it is not a simple solution. To complicate matters, the solution contains functions defined in terms of integral solutions to differential equations, which cannot be expressed as elementary functions. Needles to say, this makes getting a third term in the asymptotic expansion difficult. Let alone a fourth, and so on.

There is an alternative approach to completing this intermediate matching procedure which does not require the exact solution for the B term (or C term). It is a shortcut. However before detailing how this shortcut works, it is necessary to understand what the next step in the intermediate matching procedure is, after the B (and maybe C) term has been obtained. So, for the moment assume that they have been obtained. The intermediate approximation can then be written as

$$U_\pm=\exp\left(\pm\frac{8}{3}r^{3/2}+B_\pm+C_\pm\right)$$

where B and C are known functions of r and ϵ and in this case might contain constants n and λ. The subscript $\pm$ has been added to indicate that this expression contains two independent approximations to two independent solutions of the original differential equation (as was the case with the global $\epsilon\to 0$ approximation.).

With this intermediate approximation at hand, the matching is done as follows:

- First, take the $r\to 0$ limit of the intermediate approximation $U_\pm$.
- Then, make the same stretching transformation (with the chosen distinguished limit)

$$s=\frac{1}{2}-\epsilon^{2/3}r$$

to the approximation that is being matched on the "left end" of the r domain (that is r close to 0). In this case, this is the S_0 and $S_{3/2}$ solutions obtained in the $s \to 1/2$ limit[1].

- Take the $\epsilon \to 0$ limit of this result.
- If there is a domain of overlap, it should now be possible to write the $U_\pm$ approximations as linear combinations of the S_0 and $S_{3/2}$ approximations, in terms of the above expansions.
- This process is then repeated on the other side of the r domain. Take the $r \to \infty$ limit of $U_\pm$. Perform the stretching transformation to the $S_\pm$ approximation. Take the $\epsilon \to 0$ limit of that result (which basically forces $s \to 1/2$), and then match.

If the above process is successful, the result would be an equation of the form

$$S_\pm = C1_\pm U_+ + C2_\pm U_-$$

and another of the form

$$U_\pm = C3_\pm S_0 + C4_\pm S_{3/2},$$

where $C1_\pm, ..., C4_\pm$ are eight constants determined by the matching procedure above. From these two relationships, the link between the global $S_\pm$ approximation and the local $S_{0,3/2}$ approximation can be obtained.

Now, the salient question is, are the $B_\pm$ and $C_\pm$ terms really necessary for performing the procedure delineated above? Is it not enough to have the leading term

$$A_\pm = \pm\frac{8}{3}r^{3/2}$$

already obtained?

On the face of it, the answer is yes, typically the B term is necessary. As was discussed in the paragraph leading up to, and after equation (9.3), just because the arguments of exponentials (call them $Arg1$ and $Arg2$) satisfy

$$Arg1 \sim Arg2$$

[1]The fact that the $\to 0$ limit is matched to the $s - 1/2$ limit seems backwards at first, but is clear considering stretching transformation $r = \epsilon^{-2/3}(1/2 - s)$. The limit $r \to 0$ means that s is going to $1/2$ faster than ϵ goes to zero. The reverse in order comes from the fact that in the transformation, as s increases, r decreases.

it does not mean that

$$\exp(Arg1) \sim \exp(Arg2)$$

(in the appropriate limit). To guarantee this, the asymptotic approximation of the arguments must be carried out far enough to where the difference between them is much less than O(1) (in the asymptotic limit).

Now, in the particular situation at hand, the A term is of O(1) as $\epsilon \to 0$, which means that the B term must be $<< O(1)$. So on that account, the A term could be enough. This is often not the case when dealing with singular asymptotics, as will be evidenced in ch. 11. Also, even in the case at hand, using a single term to match series is not very confidence inspiring, so it is still a good idea to get at least one more term in the asymptotic approximation.

As far as the alternative technique goes, the key is to realize that while getting the B term (and more) is often necessary, the exact value is not needed. Only the the asymptotic behavior in the limits $r \to 0$ and $r \to \infty$ are actually needed, since the matching will be done in those limits. Moreover, these limits can be obtained straight from the equation governing B using the same techniques already employed above. It turns out that the asymptotic behavior for this B term can be obtained to as many orders as desired.

At first glance it may seem like this trick can only work once. There is no need to get the exact B term since the asymptotic limits can be obtained from the equation that defines it. However, without the exact B term to plug back in, there is no way to obtain the equation for the next order term - the C term. So its asymptotic behaviors cannot be obtained in the same way. However, this is not the case. Using educated guesses as to the completed structure of B (that are later corroborated), the equation for C can be obtained and its asymptotic limits derived. The detailed use of this shortcut for the problem at hand is relegated to Appendix C. Its results are as follows.

$r \to \infty$ asymptotics

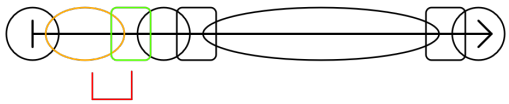
$$U_{\pm} = \exp\left(\pm\frac{8}{3}r^{3/2} + \left(\pm\frac{12}{5}r^{5/2} + (n-2)r \mp \frac{(n-2)}{8}r^{-1/2}\right)\epsilon^{2/3} \pm \frac{23}{7}r^{7/2}\epsilon^{4/3}\right). \tag{9.8}$$

$r \to 0$ asymptotic

$$U_\pm = \exp\left(\pm \frac{8}{3} r^{3/2} + \left(\pm \frac{4}{5}(4n-5) r^{5/2} - \frac{32}{5}(n-2) r^4 \right) \epsilon^{2/3} \right.$$
$$\left. \pm \frac{1}{7}\left(16n^2 - 24n + 7\right) r^{7/2} \epsilon^{4/3} + O(\epsilon)^2 \right) \tag{9.9}$$

It is important to note that these are simply two different limits of the same approximate solutions $U_\pm$; namely the intermediate approximation of mixed local and global limits. That is the U_+ in equation (9.8) is the $r \to \infty$ limit of a function whose $r \to 0$ limit is the U_+ in equation (9.9). And the same goes for the U_-.

Matching Intermediate with Global $\epsilon \to 0$



To match the first part, take the global $\epsilon \to 0$ approximation (9.4) (using only the A,B and C terms) and perform the stretching transformation used to get to the intermediate approximation

$$s = \frac{1}{2} - r\epsilon^{2/3} \tag{9.10}$$

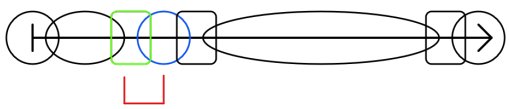
and then expanding the argument about $\epsilon \to 0$ yields

$$S_\pm \sim \exp\left(\frac{n-2}{2} \ln(2) \mp \frac{8}{3} r^{3/2} + \left(\mp \frac{12}{5} r^{5/2} + (n-2) r \pm \frac{(n-2)}{8} r^{-1/2} \right) \epsilon^{2/3} \right.$$
$$\left. + \left(\mp \frac{23}{7} r^{7/2} + (n-2) r^2 \mp \frac{1}{16}\left(4n^2 - 17n + 18\right) r^{1/2} \right) \epsilon^{4/3} \right) \tag{9.11}$$

So the matching matching between the global approximation and the $r \to \infty$ limit of the intermediate approximation is:

$$S_\pm = s^{(n-2)/2} U_\mp \tag{9.12}$$

Matching Intermediate to local $s \to 1/2$



Similarly at the other end, to match the $r \to 0$ expansion of the intermediate approximation to the local $s \to 1/2$ approximation, start with the local solutions $y_0(s)$ and $y_{3/2}(s)$ (see equation 9.2), plug in the stretching transformation

$$s = \frac{1}{2} - r\epsilon^{2/3} \tag{9.13}$$

and expand about $\epsilon \to 0$. This yields[2]

$$\begin{aligned} y_{3/2}(r) &\sim i\epsilon\left(r^{3/2} + \frac{3}{10}(4n-5)r^{5/2}\epsilon^{2/3} + \frac{3}{56}\left(16n^2 - 24n + 7\right)r^{7/2}\epsilon^{4/3} + O(\epsilon)^2\right) \\ y_0(r) &\sim 1 + \frac{32}{9}r^3 + \lambda^2 r^3\epsilon^2 + O(\epsilon)^{8/3}. \end{aligned} \tag{9.14}$$

On the other hand, expanding the exponential in the $r \to 0$ approximation (eq 9.9) about $r \to 0$ yields

$$U_\pm = 1 \pm \frac{8}{3}r^{3/2} \pm \frac{4}{5}(4n-5)r^{5/2}\epsilon^{2/3} + \frac{32}{9}r^3 \pm \left(16n^2 - 24n + 7\right)r^{7/2}\epsilon^{4/3} + \ldots \tag{9.15}$$

So the matching between local $s \to 1/2$ and the $r \to 0$ limit of the intermediate approximation is

$$U_\pm = y_0 \mp \frac{i}{\epsilon}\frac{8}{3}y_{3/2}. \tag{9.16}$$

Putting these two matching results together leads to the following match between the global $\epsilon \to 0$ approximation and the local $s \to 1/2$ approximation:

$$S_\pm(s) = 2^{(n-2)/2}\left(y_0 \pm \frac{i}{\epsilon}\frac{8}{3}y_{3/2}\right) \tag{9.17}$$

[2]The same choice of root was made as the choice made in the paragraph before equation 6.1. Namely $(-r\epsilon^{2/3})^{1/2} = ir^{1/2}\epsilon^{1/3}$

9.5 Collected Results

The local $s \to 0$ solution that doesn't blow up at the origin is a constant multiple of the global (approximate) solution $S_+(s)$. On the other side, matching the global $\epsilon \to 0$ to the local $s \to 1/2$ via an intermediate approximation resulted in

$$S_\pm(s) = 2^{(n-2)/2}\left(y_0 \pm \frac{i}{\epsilon}\frac{8}{3}y_{3/2}\right). \tag{9.18}$$

Put together this means that the solution that is regular at the origin is proportional to

$$y_0 + \frac{i}{\epsilon}\frac{8}{3}y_{3/2} \tag{9.19}$$

This is precisely the result obtained in chapter 6 for the case of $n = 2$ and $\lambda = 0$ (see equation 6.1 - replacing k with ϵ^{-1}). Therefore as foreseen, the matching is not dependent on dimension or mass.

One important conclusion is that there is no new information about geodesics that can be gleaned from this higher dimensional, general mass scenario. The comments made in section 6.4, about how the Klein- Gordon geometric optics limit provides no real guidance in understanding the behavior of particles across the transition surface, still hold. However, with the work done in this chapter, it becomes clear why the geometric optics limit leads to no new understanding about the behavior of geodesics across the transition. Namely, the geometric optics approximation breaks down near the transition surface. This is evidenced by the fact that the global $\epsilon \to 0$ approximation (geometric optics) could not be matched directly to the local approximation at the transition ($s \to 1/2$). The geometric optics dominant balance leaves out terms that become more and more physically relevant as the transition is approached. Technically, the matching was performed in the Riemannian region. However, it should come as no surprise that the exact same behavior occurs when attempting to match a global $\epsilon \to 0$ approximation in the Lorentzian domain with the $s \to 1/2$ approximation.

One way to interpret this is that the particle interpretation of matter breaks down near the transition, therefore to speak of particles is contextually wrong. In the regime near the transition, matter is fundamentally a wave. So in keeping with this new understanding about the fundamental wave-like character of matter, I turn to an exploration of the quantum behavior of the Klein-Gordon equation. In particular,

I consider the question of particle creation in a Quantum Field theory in a (fixed) curved spacetime. Chapter 10 provides a lightning introduction to quantum field theory on curved spacetimes (QFTCS), and then chapter 11 investigates the question of particle production in the context the transition universe.

Chapter Ten: Interlude - QFTCS and Particle Creation

This chapter presents a minimal introduction to quantum field theory in curved spacetime. A Klein-Gordon field is treated quantum mechanically and spacetime treated as a classical, fixed background. The approach is that of canonical (Hamiltonian) quantization starting from an action. A quantum harmonic oscillator is used as a blueprint for constructing the Hilbert Space of states. Vacuum ambiguity and particle creation are discussed.

10.1 Classical

In this introduction, I will work with the simplest case applicable to the problem under investigation in this thesis, leaving many complications outside the scope of this chapter. For more complete presentations of this framework see Birrell and Davies [1984] and Mukhanov and Winitzki [2007]. In particular, I will use a Klein-Gordon field, minimally coupled to a 2D classical spacetime with a metric of the form: $ds^2 = -N(t)dt \otimes dt + h(t)dx \otimes dx$. This theory is described the action:

$$S[\psi] = \int \left(\frac{1}{2N}\dot{\psi}^2 - \frac{1}{2h}(\partial_x \psi)^2 - \frac{1}{2}\mu^2\psi^2 \right) \sqrt{Nh}\; dtdx. \tag{10.1}$$

Following the canonical approach to quantization, the first step is to obtain the Hamiltonian description of the theory. So, in the action above, the Lagriangian density is

$$\mathscr{L} = \left(\frac{1}{2N}\dot{\psi}^2 - \frac{1}{2h}(\partial_x \psi)^2 - \frac{1}{2}\mu^2\psi^2 \right) \sqrt{Nh},$$

which is used to obtain the canonically conjugate momentum to the field strength ψ, namely

$$\begin{aligned} \pi(t,x) &:= \frac{\delta \mathscr{L}}{\delta \dot{\psi}(t,x)} \\ &= \sqrt{\frac{h}{N}}\dot{\psi}(t,x). \end{aligned}$$

The Hamiltonian can then be constructed as

$$H := \int \left(\pi\dot{\psi} - \mathscr{L}\right) dx \tag{10.2}$$

$$= \int \left(\frac{1}{2h}\pi^2 + \frac{1}{2h}(\partial_x\psi)^2 + \frac{1}{2}\mu^2\psi^2\right)\sqrt{Nh}\; dx. \tag{10.3}$$

so Hamiltonian density is

$$\mathscr{H} = \left(\frac{1}{2h}\pi^2 + \frac{1}{2h}(\partial_x\psi)^2 + \frac{1}{2}\mu^2\psi^2\right)\sqrt{Nh}. \tag{10.4}$$

From the Hamiltonian, the equations of motion for ψ and π are obtained by way of the Poisson bracket:

$$\begin{aligned}
\partial_t\psi(t,x) &= \{\psi(t,x), H(t)\} \\
&:= \int d^3y \left(\frac{\delta\psi(t,x)}{\delta\psi(t,y)}\frac{\delta H(t)}{\delta\pi(t,y)} - \frac{\delta\psi(t,x)}{\delta\pi(t,y)}\frac{\delta H(t)}{\delta\psi(t,y)}\right) \\
&= \int dy\delta(x-y)\sqrt{\frac{N}{h}}\pi(t,y) \\
&= \sqrt{\frac{N}{h}}\pi(t,x)
\end{aligned}$$

and doing the same for π yields

$$\partial_t\pi(t,x) = \sqrt{\frac{N}{h}}\partial_x^2\psi - \mu^2\psi\sqrt{Nh}. \tag{10.5}$$

For completeness sake, putting these two equations together results in

$$\partial_t^2\psi + \frac{N}{2h}\frac{d}{dt}\left(\frac{h}{N}\right)\partial_t\psi - \frac{N}{h}\partial_x^2\psi + N\mu^2\psi = 0. \tag{10.6}$$

Because of the isotropy of space, a Fourier decomposition of the fields leads to a very useful decoupling of degrees of freedom. Letting

$$\psi(t,x) = \int \frac{dk}{\sqrt{2\pi}} e^{ikx}\psi_k(t)$$

and plugging it into equation (10.6), implies

$$\ddot{\psi}_k + \frac{N}{2h}\frac{d}{dt}\left(\frac{h}{N}\right)\dot{\psi}_k + \frac{N}{h}\left(k^2 + h\mu^2\right)\psi_k = 0, \tag{10.7}$$

so each Fourier mode is independent of all the other Fourier modes. The Hamiltonian equations of motion also decouple, viz.,

$$\begin{aligned} \dot{\psi}_k &= \sqrt{\frac{N}{h}}\pi_k \\ \dot{\pi}_k = &= -\sqrt{\frac{N}{h}}\left(k^2 + h\mu^2\right)\psi_k. \end{aligned} \tag{10.8}$$

This decoupling simplifies the construction of a quantum field theory.

10.2 Transition to Quantum

The transition from a classical field theory to a quantum field theory is accomplished, in an extremely pragmatic approach, by

- replacing the *real-valued* field strength variables $\psi(t,x), \pi(t,x)$ with *hermetian* operators $\hat{\psi}(t,x), \hat{\pi}(t,x)$ (the space on which they act is as yet undetermined),
- imposing equal time canonical commutation relations

$$[\hat{\psi}(t,x), \hat{\pi}(t,y)] = i\hbar\delta(x-y), \quad [\hat{\psi}(t,x), \hat{\psi}(t,y)] = [\hat{\pi}(t,x), \hat{\pi}(t,y)] = 0 \tag{10.9}$$

- and requiring that the operators $\hat{\psi}, \hat{\pi}$ satisfy Hamilton's equations of motion as derived in the classical theory, in this case

$$\begin{aligned} \partial_t\hat{\psi} &= \sqrt{\frac{N}{h}}\hat{\pi} \\ \partial_t\hat{\pi} &= \sqrt{\frac{N}{h}}\partial_x^2\hat{\psi} - \mu^2\hat{\psi}\sqrt{Nh}. \end{aligned}$$

These equations can also be arrived at by constructing the Hamiltonian operator $\hat{H}$, from $\hat{\psi}$ and $\hat{\pi}$, in the same way that the classical counterparts are related[1], and then replacing the Poisson bracket with the commutator, that is[2]:

$$\{,\} \to \frac{1}{i\hbar}[,].$$

[1] Up to potential ordering ambiguities that can in principle be resolved by experimentation

[2] This is tantamount to imposing the canonical commutation relations.

Notice that the metric, here determined by the functions $N(t)$ and $h(t)$ is left as a regular, classical metric. It is not considered a dynamical variable in the theory, but simply a predetermined background.

It is safe to assume that whatever concrete representation of the operators $\hat{\psi}$ and $\hat{\pi}$, as acting on a Hilbert space of states, the operation of Fourier decomposition can be well-defined. Therefore, the operators can be written as

$$\begin{aligned}\hat{\psi}(t,x) &= \int \frac{dk}{\sqrt{2\pi}} e^{ikx}\hat{\psi}_k(t) \\ \hat{\pi}(t,x) &= \int \frac{dk}{\sqrt{2\pi}} e^{ikx}\hat{\pi}_k(t).\end{aligned}$$

Since $\hat{\psi}$ and $\hat{\pi}$ are hermetian operators, the Fourier mode counterparts satisfy

$$\hat{\psi}_k^\dagger = \hat{\psi}_{-k}, \quad \hat{\pi}_k^\dagger = \hat{\pi}_{-k}, \tag{10.10}$$

and by plugging in the Fourier expansion into the canonical commutation relations (10.9), it can be shown that

$$[\hat{\psi}_k(t), \hat{\pi}_{k'}(t)] = i\hbar\delta(k+k'), \quad [\hat{\psi}_k(t), \hat{\psi}_{k'}(t)] = [\hat{\pi}_k(t), \hat{\pi}_{k'}(t)] = 0. \tag{10.11}$$

The Hamiltonian operator becomes[3]

$$\begin{aligned}\hat{H}(t) &:= \int \left(\frac{1}{2h}\hat{\pi}^\dagger(t,x)\hat{\pi}(t,x) + \frac{1}{2h}\partial_x\hat{\psi}^\dagger(t,x)\partial_x\hat{\psi}(t,x) + \frac{1}{2}\mu^2\hat{\psi}^\dagger(t,x)\hat{\psi}(t,x) \right) \sqrt{Nh}\, dx \\ &= \int \left(\frac{1}{2}\hat{\pi}_k^\dagger\hat{\pi}_k + \frac{1}{2}\left(k^2+\mu^2 h\right)\hat{\psi}_k^\dagger\hat{\psi}_k \right) \sqrt{\frac{N}{h}}\, dk,\end{aligned} \tag{10.12}$$

and the equations of motion are

$$\begin{aligned}\frac{d}{dt}\hat{\psi}_k(t) &= \frac{1}{i\hbar}[\hat{\psi}_k(t), \hat{H}(t)] \\ &= \sqrt{\frac{N}{h}}\hat{\pi}_k(t) \\ \frac{d}{dt}\hat{\pi}_k(t) &= \frac{1}{i\hbar}[\hat{\pi}_k(t), \hat{H}(t)] \\ &= -\sqrt{\frac{N}{h}}\left(k^2+\mu^2 h\right)\hat{\psi}_k(t).\end{aligned} \tag{10.13}$$

[3]Using the Hermiticity of $\hat{\psi}$, the term ψ^2, which becomes $\hat{\psi}^2$, has been written as $\hat{\psi}^\dagger\hat{\psi}$; same for π. This is just a calculation convenience.

Combining the equations of motion for $\hat{\psi}$ and $\hat{\pi}$ yields

$$\frac{d^2}{dt^2}\hat{\psi}_k(t) + \frac{N}{2h}\frac{d}{dt}\left(\frac{h}{N}\right)\frac{d}{dt}\hat{\psi}_k + \frac{N}{h}\left(k^2 + \mu^2 h\right)\hat{\psi}_k(t) = 0. \tag{10.14}$$

So, as with the classical theory, expressing the dynamical variables in terms of Fourier modes leads to a decoupling of degrees of freedom. It almost becomes a collection of quantum mechanical systems with "position" operator $\hat{\psi}_k$ and "momentum" operator $\hat{\pi}_k$, with Hamiltonian

$$\hat{H}_k(t) = \left(\frac{1}{2}\hat{\pi}_k^\dagger\hat{\pi}_k + \frac{1}{2}\left(k^2 + \mu^2 h\right)\hat{\psi}_k^\dagger\hat{\psi}_k\right)\sqrt{\frac{N}{h}}.$$

However, the fact that $\hat{\psi}_k$ and $\hat{\pi}_k$ are not Hermitian, but instead satisfy

$$\hat{\psi}_k^\dagger = \hat{\psi}_{-k}, \quad \hat{\pi}_k^\dagger = \hat{\pi}_{-k}, \tag{10.15}$$

and the commutation relations are

$$[\hat{\psi}_k(t), \hat{\pi}_{k'}(t)] = i\hbar\delta(k + k')$$

instead of the canonical ones, means that the modes k and $-k$ are coupled. In fact the conjugate momentum associated with mode $\hat{\psi}_k$ is $\hat{\pi}_k^\dagger$, as evidenced by the commutation relations. Alternatively, this can also be seen by expressing the action in terms of the Fourier modes. This does not present a serious challenge, since the theory can be rewritten in terms of Hermitian operators by splitting $\hat{\psi}_k$ into $\hat{\psi}_k^{(1)} + i\hat{\psi}_k^{(2)}$, with constraints imposed by (10.15). However, it is more convenient to keep as is; this quirk will not obstruct the construction of the Hilbert space in the next section.

10.3 Hilbert Space of States

Notice that if $N = h = 1$, the Fourier modes almost[4] reduce to a bunch of decoupled harmonic oscillators. This similarity is exploited to construct the Hilbert space of states in much the same way it is typically constructed for a quantum harmonic oscillator (QHO).

[4] "Almost", since as discussed above some modes are coupled due to the equation (10.15).

QHO recipe

A simple quantum harmonic oscillator with mass set to unity, is described by position and momentum operators $\hat{q}$ and $\hat{p}$, and Hamiltonian

$$\hat{H} = \frac{1}{2}\hat{p}^2 + \frac{1}{2}\omega^2\hat{q}^2.$$

In the Heisenberg picture of time evolution, the operators $\hat{q} = \hat{q}(t)$ and $\hat{p} = \hat{p}(t)$ are time dependent, and their evolution equations are

$$\begin{aligned} \frac{d}{dt}\hat{q}(t) &= \frac{1}{i\hbar}[\hat{q}(t), \hat{H}] \\ &= \hat{p}(t) \\ \frac{d}{dt}\hat{p}(t) &= \frac{1}{i\hbar}[\hat{p}(t), \hat{H}] \\ &= -\omega^2\hat{q}(t), \end{aligned}$$

which when combined, yield

$$\frac{d^2}{dt^2}\hat{q}(t) = -\omega^2\hat{q}(t). \tag{10.16}$$

Often the first step towards "solving" the QHO is to introduce creation and annihilation operators

$$\begin{aligned} \hat{a}^- &:= \sqrt{\frac{\omega}{2\hbar}}\left(\hat{q} + \frac{i}{\omega}\hat{p}\right) \\ \hat{a}^+ &:= \left(\hat{a}^-\right)^\dagger = \sqrt{\frac{\omega}{2\hbar}}\left(\hat{q} - \frac{i}{\omega}\hat{p}\right). \end{aligned}$$

They satisfy the equal-time commutation relations

$$[\hat{a}^-(t), \hat{a}^+(t)] = 1, \quad [\hat{a}^-(t), \hat{a}^-(t)] = [\hat{a}^+(t), \hat{a}^+(t)] = 0, \tag{10.17}$$

and can be used to express the Hamiltonian as

$$\hat{H} = \hbar\omega\left(\hat{a}^+\hat{a}^- + \frac{1}{2}\right). \tag{10.18}$$

This in turn can be used to derive the equations of motion

$$\frac{d}{dt}\hat{a}^\pm = \pm i\omega\hat{a}^\pm \tag{10.19}$$

whose solutions are

$$\hat{a}^\pm(t) = \hat{a}_0^\pm e^{\pm i\omega t} \tag{10.20}$$

where $\hat{a}_0^{\pm}$ are time independent operators[5]. They inherit the commutation relations

$$[\hat{a}_0^-, \hat{a}_0^+] = 1, \quad [\hat{a}_0^-, \hat{a}_0^-] = [\hat{a}_0^+, \hat{a}_0^+] = 0.$$

These time independent creation and annihilation operators can now be used to retroactively construct the Hilbert space which represent the space of states.

First, note that the Hamiltonian can be rewritten using the time-independent ladder operators as follows

$$\hat{H} = \hbar\omega \left(\hat{a}_0^+ \hat{a}_0^- + \frac{1}{2} \right), \tag{10.21}$$

and the commutation relations between the Hamiltonian and the ladder operators are

$$[\hat{H}, \hat{a}_0^+] = \hbar\omega \hat{a}_0^+ \quad [\hat{H}, \hat{a}_0^-] = -\hbar\omega \hat{a}_0^-. \tag{10.22}$$

Next, the existence of a state $|0\rangle$ is postulated[6], which satisfies

$$\hat{a}_0^- |0\rangle = 0. \tag{10.23}$$

This is state is an eigenstate of the Hamiltonian with eigenvalue $1/2\hbar\omega$. From this state, a basis of states for the Hilbert space is constructed via repeated applications of the creation operator $\hat{a}_0^+$. The resulting vectors are all eigenvectors of the Hamiltonian operator. They are

$$|n\rangle := \frac{1}{\sqrt{n!}} \left(\hat{a}_0^+ \right)^n |0\rangle. \tag{10.24}$$

Once the Hilbert space is defined, the action of position and momentum operators on the states can be obtained by expressing them in terms of the $\hat{a}_0^{\pm}$. For example:

$$\hat{q}(t) = \frac{1}{\sqrt{2}} \left(\sqrt{\frac{\hbar}{\omega}} e^{-i\omega t} \hat{a}_0^- + \sqrt{\frac{\hbar}{\omega}} e^{i\omega t} \hat{a}_0^+ \right). \tag{10.25}$$

It is this expression of $\hat{q}$ in terms of creation and annihilation operators that is used as a a point of departure for building the QFT Hilbert space.

[5]Basically, they are constants of integration, which represent the value of $\hat{a}^{\pm}$ at $t = 0$.

[6]A slightly more rigorous approach starts by positing the existence of an eigenvector of the Hamiltonian and then, using the commutation relations with the annihilation operator and the positive definiteness of the Hamiltonian, obtain a state which must be annihilated by $\hat{a}$.

Mode Expansion

Returning to the quantum field theory problem at hand, the construction of the Hilbert space of states begins with a generalization of equation (10.25), namely, letting

$$\hat{\psi}_k(t) = \frac{1}{\sqrt{2}} \left(\hat{a}_k^- v_k^*(t) + \hat{a}_{-k}^+ v_k(t) \right) \tag{10.26}$$

where $v_k(t)$ is a complex valued function of time, and the $\hat{a}_k^\pm$ are already the time-independent ladder operators. The subscript 0 has been dropped since it makes the notation simpler, and the time dependence has already been isolated from the ladder operators in the ansatz above.

There are a couple of important points to make about this ansatz:

- Since all the time dependence of $\hat{\psi}_k$ has been placed in the functions $v_k(t)$ and $v_k^*(t)$, if $v_k(t)$ satisfies the equation of motion, namely

$$\ddot{v}_k + \frac{N}{2h} \frac{d}{dt} \left(\frac{h}{N} \right) \dot{v}_k + \frac{N}{h} \left(k^2 + h\mu^2 \right) v_k = 0, \tag{10.27}$$

  then $\hat{\psi}_k(t)$ will satisfy the operator version of the equation of motion (i.e. equation 10.14). The functions $\{v_k(t)\}$ are called the **mode functions**, the ansatz (10.26) is called a **mode expansion**, and the equation of motion takes on the name **mode equation**.

- Since the *mode equation* depends only on the magnitude of k (because of the isotropy of space) the mode functions can be chosen in such a way that $v_k(t) = v_{-k}(t)$, for all k. Under this assumption, the ansatz (10.26) satisfies the requirement that

$$\hat{\psi}_k^\dagger(t) = \hat{\psi}_{-k}(t). \tag{10.28}$$

- Finally, a necessary condition for the operators $\hat{a}_k^\pm$ to act as ladder operators is that they satisfy the commutation relations

$$[\hat{a}_k^-, \hat{a}_{k'}^+] = \delta(k - k'), \qquad [\hat{a}_k^-, \hat{a}_{k'}^-] = [\hat{a}_k^+, \hat{a}_{k'}^+] = 0 \tag{10.29}$$

These are simply a generalization of the QHO ladder commutation relations to the case of uncountably many degrees of freedom.[7] Consistency between these commutation relations and the commutation relations

$$[\hat{\psi}_k(t), \hat{\pi}_{k'}(t)] = i\hbar\delta(k + k')$$

leads to the condition

$$(v_k\dot{v}_k^* - v_k^*\dot{v}_k) = -2i\hbar\sqrt{\frac{N}{h}}.$$

To see this, first note that the ansatz (10.26) implies that

$$\hat{\pi}_k(t) = \sqrt{\frac{h}{2N}}\left(\hat{a}_k^-\dot{v}_k^*(t) + \hat{a}_{-k}^+\dot{v}_k(t)\right).$$

Substituting the mode expansions into the commutator of $\hat{\psi}_k$ and $\hat{\pi}_{k'}$ yields

$$\begin{aligned}[\hat{\psi}_k, \hat{\pi}_{k'}] &= \sqrt{\frac{h}{4N}}\Bigg(v_k^*\dot{v}_{k'}[\hat{a}_k^-, \hat{a}_{-k'}^+] + v_k\dot{v}_{k'}^*[\hat{a}_{-k}^+, \hat{a}_{k'}^-] \\ &\quad + v_k^*\dot{v}_k^*[\hat{a}_k^-, \hat{a}_{k'}^-] + v_k\dot{v}_{k'}[\hat{a}_{-k}^+, \hat{a}_{-k'}^+]\Bigg).\end{aligned}$$

Imposing the commutation relations of the Fourier modes and the desired commutation relations for the ladder operators results in

$$i\hbar\delta(k + k') = \sqrt{\frac{h}{4N}}\left(v_k^*\dot{v}_k - v_k\dot{v}_k^*\right)\delta(k + k').$$

This in turn implies

$$(v_k\dot{v}_k^* - v_k^*\dot{v}_k) = -2i\hbar\sqrt{\frac{N}{h}} \tag{10.30}$$

which is the aforementioned condition. Note that the expression on the left hand side of (10.30) is just the Wronskian of v_k and v_k^*. So it can be written as

$$W[v_k, v_k^*] = -2i\hbar\sqrt{\frac{N}{h}}. \tag{10.31}$$

[7] These uncountably many degrees of freedom actually present a serious mathematical challenge when it comes to building a Hilbert space; they lead to a non-separable Hilbert space. So typically, to avoid this, one considers a quantum field theory in a very large but finite box, with nice boundary conditions. This leads to the Fourier modes becoming countable and manageable. As always, the ultimate arbiter of truth is experimentation. So far, this procedure, together with the resulting Hilbert space, seem to work just fine.

It turns out that this amounts to a simple normalization condition. This can be seen by deriving the first order differential equation satisfied by the Wronskian. Taking a time derivative of $W[v_k, v_k^*]$ and then using the mode equation (10.27) to replace second derivatives of v_k and v_k^*, yields

$$\frac{d}{dt}W[v_k, v_k^*] = W[v_k, v_k^*]\left(-\frac{N}{2h}\frac{d}{dt}\left(\frac{h}{N}\right)\right).$$

The solution to this equation is

$$W[v_k, v_k^*] = C\sqrt{\frac{N}{h}} \tag{10.32}$$

where C is a constant of integration. By comparing the general Wronskian solution (10.32) to the condition (10.31), it can be seen that the condition imposed by requiring consistency between commutation relations amounts to a simple normalization of the mode functions. That is, given any choice of mode functions, this consistency condition can be achieved using a new mode equation that is a constant multiple of the old one. From here on out, this consistency condition will be called the **Wronskian condition**.

Plugging in the mode expansion into the Hamiltonian in eq (10.12), yields (after some simplification)

$$\hat{H} = \sqrt{\frac{N}{h}}\int\left(\left(\hat{a}_k^+\hat{a}_k^- + \frac{1}{2}\right)E_k(t) + \frac{1}{2}\hat{a}_k^+\hat{a}_{-k}^+F_k(t) + \frac{1}{2}\hat{a}_{-k}^-\hat{a}_k^-F_k^*(t)\right)dk$$

where

$$\begin{aligned} E_k &:= \frac{h}{2N}|\dot{v}_k|^2 + \frac{1}{2}(k^2 + h\mu^2)|v_k|^2 \\ F_k &:= \frac{h}{2N}(\dot{v}_k)^2 + \frac{1}{2}(k^2 + h\mu^2)(v_k)^2. \end{aligned} \tag{10.33}$$

The commutation relations between the ladder operators and the Hamiltonian operator are

$$\begin{aligned} \left[\hat{H}, \hat{a}\right] &= -\sqrt{\frac{N}{h}}\left(E_k\hat{a}_k^- + F_k\hat{a}_{-k}^+\right) \\ \left[\hat{H}, \hat{a}_k^+\right] &= \sqrt{\frac{N}{h}}\left(E_k\hat{a}_k^+ + F_k^*\hat{a}_{-k}^-\right). \end{aligned} \tag{10.34}$$

These are not quite analogous to the QHO ladder operator commutation relations. In particular, for the QHO, if $|\lambda\rangle$ is an eigenvector of the Hamiltonian with eigenvalue λ, then the commutation relations (10.22) guarantee that $\hat{a}_0^+|\lambda\rangle$ is also an eigenvector of the Hamiltonian, with eigenvalue $\lambda + \hbar\omega$. This is not the case for the QFT generalization above. The existence of F_k spoils this. I will table this point for the moment, and continue as if F_k where equal to 0. The resolution to this problem will come as a consequence of a resolution to another problem, namely the ambiguity in the choice of mode functions, which I will address momentarily. So, assuming for the time being that $F_k = 0$, the commutation relations of the ladder operators with the Hamiltonian allow for a Hilbert space construction almost entirely analogous to the (multiple) QHO problem. The existence of a state $|0\rangle$ is postulated which has the property that

$$\hat{a}_k^-|0\rangle = 0$$

for all k. This state is an eigenstate of the Hamiltonian with eigenvalue[8]

$$\sqrt{\frac{N}{h}} \int \frac{1}{2} E_k(t) dk. \tag{10.35}$$

and is called the **minimum excitation** state. Furthermore, excited states are defined as

$$|\{n_k\}\rangle := \prod_k \frac{(\hat{a}_k^+)^{n_k}}{\sqrt{n_k!}}|0\rangle, \tag{10.36}$$

where the $\{n_k\}$ are the **occupation numbers**, which can be interpreted as specifying the number of particles in each mode k. These are all eigenstates of the Hamiltonian operator. And thus the Hilbert is constructed as the space of all (mathematically appropriate) linear combinations of these energy eigenstates.

10.4 Ambiguity in the Vacuum

There is an ambiguity in the definition of the *minimum excitation* state $|0\rangle$. It stems from the choice of ansatz (10.26). In writing the expansion of $\hat{\psi}_k$ in terms of $\hat{a}_k^-$ and $\hat{a}_k^+$, a concrete choice of *mode function* v_k must be made. If a different choice were made, say $w_k \neq v_k$, the ladder operators would be different, and therefore the state annihilated by all the ladder operators would also be different.

[8]Generally, this is a divergent quantity that needs to be regularized in some way.

The details of this argument go as follows. Since the mode function is a second order, linear, homogeneous ordinary differential equation, and the solutions v_k and v_k^* are linearly independent (by virtue of non-vanishing Wronskian), w_k can be written as

$$w_k = \alpha_k v_k + \beta_k v_k^* \tag{10.37}$$

for some fixed coefficients $\{\alpha_k, \beta_k\} \subset \mathbb{C}$. However, since w_k must also satisfy the Wronskian condition (for the ladder operator interpretation to work), the choice of coefficients is not arbitrary. Writing the left hand side Wronskian condition (10.30) for w_k and expanding w_k in terms of v_k, yields

$$w_k \dot{w}_k^* - w_k^* \dot{w}_k = \left(|\alpha_k|^2 - |\beta_k|^2\right)\left(v_k \dot{v}_k^* - v_k^* \dot{v}_k\right). \tag{10.38}$$

This shows that if v_k satisfies the Wronskian condition, then for w_k to satisfy it also, the coefficients α_k, β_k must satisfy

$$|\alpha_k|^2 - |\beta_k|^2 = 1 \tag{10.39}$$

for all k. A transformation of the form (10.37) subject to the condition (10.39) is known as a **Bogolyubov transformation**.

Now, given these two distinct choices of mode functions, the operator $\hat{\psi}_k$ can be expanded using the mode functions v_k

$$\hat{\psi}_k(t) = \frac{1}{\sqrt{2}}\left(\hat{a}_k^- v_k^*(t) + \hat{a}_{-k}^+ v_k(t)\right)$$

or using the mode function w_k

$$\hat{\psi}_k(t) = \frac{1}{\sqrt{2}}\left(\hat{b}_k^- w_k^*(t) + \hat{b}_{-k}^+ w_k(t)\right).$$

Setting these two expansions equal to each other (since it is one and the same operator $\hat{\psi}_k$) and using the Bogolyubov transformation yields the relationships

$$\hat{a}_k^- = \alpha_k^* \hat{b}_k^- + \beta_k \hat{b}_{-k}^+, \quad \hat{a}_k^+ = \alpha_k \hat{b}_k^+ + \beta_k^* \hat{b}_{-k}^-. \tag{10.40}$$

Finally, note that the state that is annihilated by all the $\hat{a}_k$ operators, is, in general, not annihilated by all the $\hat{b}_k$ operators. That is

$$\hat{a}_k^- |0\rangle = 0 \tag{10.41}$$

implies

$$\hat{b}_k^-|0\rangle = -\frac{\beta_k}{\alpha_k^*}\left(\hat{b}_{-k}^+|0\rangle\right) \tag{10.42}$$

and similarly the other way around: The state annihilated by all the $\hat{b}_k^-$ is in general not annihilated by the $\hat{a}_k^-$. What this means is is that the minimum excitation state postulated when using one mode expansion is different than the minimum excitation state postulated when using a different mode expansion. Because of this, it is notationally appropriate to distinguish these states. From now on, when not clear from context, the minimum excitation state that results from choosing an expansion in terms of mode functions v_k will be denoted $|0\rangle_v$.

A consequence of this difference is that while the state $|0\rangle_w$ has no particles of type $\hat{b}_k^+$, it is not devoid of particles of type $\hat{a}_k^+$. In fact, the number of $\hat{a}_k^+$ particles is

$$\begin{aligned}
{}_w\langle 0|\hat{a}_k^+\hat{a}_k|0\rangle_w &= {}_w\langle 0|\left(\alpha_k\hat{b}_k^+ + \beta_k^*\hat{b}_{-k}^-\right)\left(\alpha_k^*\hat{b}_k^- + \beta_k\hat{b}_{-k}^+\right)|0\rangle_w \\
&= |\beta_k|^2\,{}_w\langle 0|\hat{b}_{-k}^-\hat{b}_{-k}^+|0\rangle_w \\
&= |\beta_k|^2\,{}_w\langle 0|\hat{b}_{-k}^+\hat{b}_{-k}^- + \delta(0)|0\rangle_w \\
&= |\beta_k|^2\delta(0).
\end{aligned} \tag{10.43}$$

The $\delta(0)$ factor comes from the uncountably infinite number of Fourier modes in space and is typically tamed via some form of discretization of the Fourier modes. The important point however, is that the density of particles, for mode k is proportional to the absolute square of the Bogolyubov coefficient β_k.

10.5 Instantaneous Lowest Energy Vacuum

So, which is "the right" mode expansion to use which leads to the real vacuum and the real notion of particles? There is no universally accepted prescription for defining the "true vacuum", mostly because, as evidenced by the previous discussion, vacuum and particle are not observer/context independent notions. However, there are a couple of useful definitions whose validity depends on the problem under consideration. Here, I will consider the **instantaneous lowest energy vacuum**.

The prescription for obtaining the *instantaneous lowest energy vacuum*, as the name suggests, is to use the Hamiltonian operator to choose among all the possible

minimum excitation states. That is, choose the mode functions w_k such that

$$ {}_w\langle 0|\hat{H}|0\rangle_w \tag{10.44}$$

takes the minimum value amongst all other $|0\rangle_v$. This is a physical criterion for choosing a preferred minimum excitation state as the vacuum. However, in general, there is no choice of mode function whose associated minimum excitation state minimizes the energy for all time. The choice of mode functions that lead to a minimum energy state at one instant in time will not be the choice that leads to the minimum energy state at a later time. Hence the qualifier *instantaneous*.

Given that the Hamiltonian takes the form

$$\hat{H} = \sqrt{\frac{N}{h}} \int \left(\left(\hat{a}_k^+ \hat{a}_k^- + \frac{1}{2} \right) E_k(t) + \frac{1}{2} \hat{a}_k^+ \hat{a}_{-k}^+ F_k(t) + \frac{1}{2} \hat{a}_{-k}^- \hat{a}_k^- F_k^*(t) \right) dk$$

the minimization condition (10.44) is equivalent to minimizing

$$E_k(t) := \frac{h}{2N} |\dot{v}_k|^2 + \frac{1}{2}(k^2 + h\mu^2)|v_k|^2$$

over all (appropriately normalized) mode functions. Typically, the way this concept is applied is by obtaining the instantaneous lowest energy vacuum in some distant past, call it $|0\rangle_{in}$ and then obtaining it again for some distant future, call it $|0\rangle_{out}$, and comparing these two states (and the associated particle structure of each). The mode functions associated with these vacuum states will be related by a Bogolyubov transformation, and the number of "out"-particles in the "in"-vacuum (and vice-versa) is determined by the β_k coefficients (see eq. 10.43).

Now, there is one loose end that needs to be tied, and that is the potentially non-zero value of the function F_k. Recall that the existence of this function spoils the commutation relations (10.34). The conception of a particle, in the the current construction, hinges on the creation operators adding these fixed quanta of energy to an energy eigenvalue state. A non-vanishing F_k term destroys this behavior. However, it turns out that the mode function which minimizes the expectation value of the energy (10.44) at a given instant in time, also makes F_k vanish, at that same instant in time. Therefore, the Hamiltonian operator is diagonal in the basis of states constructed with that choice of mode function. This is a corollary of the linear algebra theorem that, given an operator A, minimizing the expectation value $\langle x|A|x\rangle$ over all normalized vectors $|x\rangle$ yields an eigenvector of A, which has the smallest eigenvalue.

Chapter Eleven: Klein-Gordon - Quantum Particle Production

This chapter contains an investigation into the particle production of a Klein-Gordon field in the transition toy model. The idea here is to consider a universe that has tunneled into existence, a-la Hawking-Hartle. Assuming that it tunnels into a state in which the Klein-Gordon is in a minimum energy state, what is the particle content observed an infinite amount of time later. The problem becomes one of matching solutions at the transition to solutions at late times. This is accomplished via the same asymptotic analysis techniques employed in chapter 9. The relevant regimes under consideration are shown in figure (11.1) below.

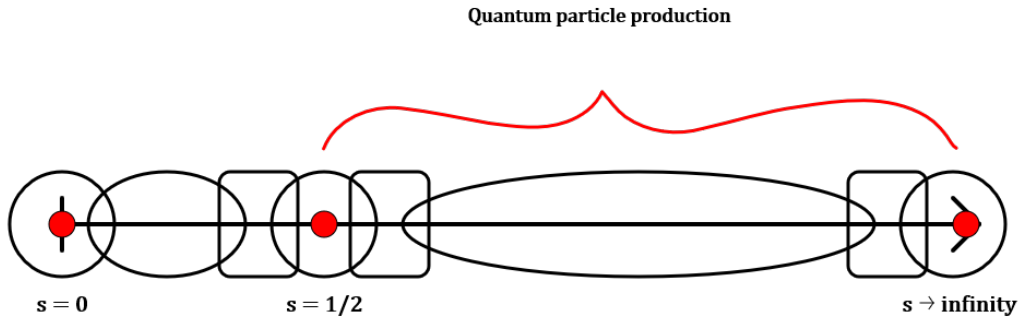


Figure 11.1: These are the regimes involved in the particle production problem.

11.1 Setup

Given the understanding developed so far, I turn to a consideration of the quantum mechanical nature of the Klein-Gordon field in a transition universe. In the Hartle-Hawking no boundary framework, the basic interpretation is that the universe tunnels into existence, out of nothing (see 2.1). The Riemannian part of spacetime is akin to the classically forbidden part of the potential in a typical tunneling problem. However, in this case it is joining emptiness on one "side" and a Lorentzian spacetime on the other. The results from the investigation performed so far suggest that right after the transition, at the temporal beginning of the Lorentzian universe, matter behaves quantum mechanically in nature. During this period, the coupling to the gravitational field is expected to produce particle excitations in the Klein-Gordon field. The question investigated in this chapter is the following: *Assuming that the universe tunnels into existence, with a Klein-Gordon field in a vacuum state, what is the particle content a t equals infinity?* To this end, I use the framework described in chapter 10 including the instantaneous lowest energy definition of vacuum. I compare the lowest energy vacuum state the just after transition to the lowest energy vacuum state as t approaches infinity.

As described in chapter 10, the starting point is the Klein-Gordon action in spacetime. The fixed background metric under consideration is

$$g = -\left(\frac{4(ct)^2}{f^2} - 1\right) d(ct) \otimes d(ct) + (ct)^2 d\theta \otimes d\theta.$$

Plugging this into the minimally coupled Klein-Gordon action

$$S_{KG} = \alpha \int \left(-\frac{1}{2} g^{\mu\nu} \partial_\mu \Psi \partial_\nu \Psi - \frac{1}{2} \frac{\mu^2 c^2}{\hbar^2} \Psi^2\right) \sqrt{|g|}\, d(ct) \wedge d\theta$$

leads to the action

$$S_{KG} = \alpha \int \mathcal{L} d(ct) \wedge d\theta$$

where the Lagrangian density is

$$\mathcal{L} = ct \sqrt{\left(\frac{4(ct)^2}{f^2} - 1\right)} \left(\frac{1}{2} \frac{(\partial_{ct}\Psi)^2}{\left(\frac{4(ct)^2}{f^2} - 1\right)} - \frac{1}{2} \frac{(\partial_\theta \Psi)^2}{(ct)^2} - \frac{1}{2} \frac{\mu^2 c^2}{\hbar^2} \Psi^2\right).$$

Once again, for simplicity, the coordinates are non-dimensionalized via the transformation $s = ct/f$. This results in the Lagrangian density

$$\mathcal{L} = s\sqrt{4s^2-1}\left(\frac{1}{2}\frac{(\partial_s\psi)^2}{(4s^2-1)} - \frac{1}{2}\frac{(\partial_\theta\psi)^2}{s^2} - \frac{1}{2}\lambda^2\psi^2\right)$$

where $\lambda = (\mu c f)/\hbar$. (Recall this is a density, so it transforms with factor of $det(J) = f$, since the coordinate transformation is $ct \to fs$.) The conjugate momentum and Hamiltonian density are

$$\pi = \frac{s}{\sqrt{4s^2-1}}\partial_s\psi \tag{11.1}$$

$$\mathscr{H} = \frac{\sqrt{4s^2-1}}{s}\left(\frac{1}{2}\pi^2 + \frac{1}{2}\left(\partial_\theta\psi\right)^2 + \frac{1}{2}\lambda^2 s^2\psi^2\right). \tag{11.2}$$

Hamilton's equations of motion are

$$\begin{aligned}\partial_s\psi(\theta) &= \frac{\sqrt{4s^2-1}}{s}\pi(\theta)\\ \partial_s\pi(\theta) &= \frac{\sqrt{4s^2-1}}{s}\left(\partial_\theta^2\psi - \lambda^2 s^2\psi\right)\end{aligned}$$

which when combined yield

$$\partial_s^2\psi - \frac{1}{s(4s^2-1)}\partial_s\psi + \frac{(4s^2-1)}{s^2}\left(-\partial_\theta^2\psi + \lambda^2 s^2\psi\right) = 0. \tag{11.3}$$

Now, in the quantization step, the fields are replaced by Hermitian operators

$$\psi(s,\theta), \pi(s,\theta) \to \hat{\psi}(s,\theta), \hat{\pi}(s,\theta)$$

with equal time commutation relations

$$\left[\hat{\psi}(s,\theta), \hat{\pi}(s,\theta')\right] = i\hbar\delta(\theta-\theta'), \quad \left[\hat{\psi}(s,\theta), \hat{\psi}(s,\theta')\right] = \left[\hat{\pi}(s,\theta), \hat{\pi}(s,\theta')\right] = 0. \tag{11.4}$$

The Hamiltonian becomes

$$\hat{H} = \int d\theta\, \frac{\sqrt{4s^2-1}}{s}\left(\frac{1}{2}\hat{\pi}^\dagger\hat{\pi} + \frac{1}{2}\partial_\theta\hat{\psi}^\dagger\partial_\theta\hat{\psi} + \frac{1}{2}\lambda^2 s^2\hat{\psi}^\dagger\hat{\psi}\right) \tag{11.5}$$

and Hamilton's equation of motion

$$\begin{aligned}\partial_s\hat{\psi} &= \frac{\sqrt{4s^2-1}}{s}\hat{\pi}\\ \partial_s\hat{\pi} &= \frac{\sqrt{4s^2-1}}{s}(\partial_\theta^2\hat{\psi} - \lambda^2 s^2\hat{\psi}).\end{aligned}$$

In chapter 10, the Fourier expansion and subsequent ladder operators expansion were presented as two separate steps. However, they can be combined into one single mode expansion step, namely, letting

$$\hat{\psi}(s,\theta) = \frac{1}{\sqrt{4\pi}} \sum_{k\in\mathbb{Z}} \left(e^{ik\theta} v_k^*(s) \hat{a}_k^- + e^{-ik\theta} v_k(s) \hat{a}_k^+ \right) \tag{11.6}$$

where $v_k(s)$ satisfies the mode equation

$$\ddot{v}_k - \frac{1}{s(4s^2-1)} \dot{v}_k + \frac{(4s^2-1)}{s^2} \left(k^2 + \lambda^2 s^2 \right) v_k = 0. \tag{11.7}$$

Note that in this case, the expansion is in terms of a sum over $k \in \mathbb{Z}$ rather than an integral over $k \in \mathbb{R}$. This is because for the problem at hand, the spatial slices are closed circles, and therefore Fourier modes are countable. The *Wronskian normalization condition* becomes

$$W[v_k, v_k^*] = -2i\hbar \frac{\sqrt{4s^2-1}}{s}. \tag{11.8}$$

Finally, plugging the mode expansion into the Hamiltonian operator yields

$$\hat{H} = \frac{\sqrt{4s^2-1}}{s} \sum_k \left(E_k(s) \left(\hat{a}_k^+ \hat{a}_k^- + \frac{1}{2} \right) + F_k(s) \frac{1}{2} \hat{a}_k^+ \hat{a}_{-k}^+ + F_k^*(s) \frac{1}{2} \hat{a}_k^- \hat{a}_{-k}^- \right) \tag{11.9}$$

where

$$E_k(s) := \frac{1}{2} \frac{s^2}{4s^2-1} |\dot{v}_k|^2 + \frac{1}{2} \omega_k^2 |v_k|^2 \tag{11.10}$$

$$F_k(s) := \frac{1}{2} \frac{s^2}{4s^2-1} (\dot{v}_k)^2 + \frac{1}{2} \omega_k^2 (v_k)^2 \tag{11.11}$$

and

$$\omega_k^2 := k^2 + \lambda^2 s^2. \tag{11.12}$$

With this setup completed, the problem now is to find the instantaneous minimum energy vacuum at two different instances in time, or in this cases two different time limits, namely $s \to 1/2$ (from above) and $s \to \infty$. Then, assuming that the universe tunnels into the vacuum state for the Klein-Gordon field - that it, the state of the Klein-Gordon field is chosen to be the instantaneous minimum energy vacuum in the limit $s \to 1/2$ - calculate the particle content with respect to the minimum energy vacuum in the limit $s \to \infty$.

Mathematically, the problem can be broken down into the following steps:

- Find a solution, v_k, to mode equation (11.7), subject to the Wronskian normalization condition (11.8), that minimizes the energy function (11.10) in the limit as s approaches $1/2$ from above.

- Repeat the above procedure but in the limit of s approaching infinity instead. Label this solution w_k.

- Express the solution w_k as a linear combination of v_k and v_k^*, to relate the ladder operators associated to the two different vacua and thus obtain the $s \to \infty$ particle count for the $s \to 1/2$ vacuum state.

Having stated the problem, there a few important points to be made about it:

- As previously discussed the notion of particles does not make much sense near the transition point, so talking about particle creation seems contradictory. However, technically, particles are not being discussed in this regime, rather only the concept of the minimum energy vacuum is being used. This is postulated as the state into which the universe tunnels. It is later, far from the transition, that particles are being considered.

- Since the mode equation (11.7) cannot be solved exactly (for general λ) asymptotic approximations in the limits s approaches $1/2$ and s approaches ∞ will be used. The energy function will be minimized in terms of these asymptotic approximations and then these will be patched together using the methods of chapter 9. It turns out that the Wronskian normalization condition and the Bogolyubov transformation property that requires the coefficients to satisfy

$$|\alpha_k|^2 - |\beta_k|^2 = 1 \tag{11.13}$$

  will be critical to the success of this method.

- The energy function that will be minimized, for each mode k, is (11.10). This ignores the leading factor of

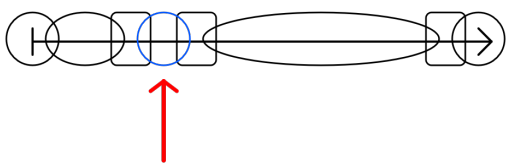
$$\frac{\sqrt{4s^2 - 1}}{s} \tag{11.14}$$

  in the Hamiltonian (11.9). Since the minimization happens for a fixed s, this factor simply contributes to an overall constant and therefore does not change the results. On the other hand, at $s = 1/2$ this factor equals zero, so it would

seem that all mode functions lead to the same energy. However, the minimization will not be done exactly at $s = 1/2$, rather it will be performed in the limit as s approaches $1/2$, in which case the leading factor is nonzero and therefore irrelevant.

11.2 Asymptotic Approximations and Energy Minimization

$s \to 1/2$ limit



To obtain approximate solutions to the mode equation

$$\ddot{v}_k - \frac{1}{s(4s^2-1)}\dot{v}_k + \frac{(4s^2-1)}{s^2}\left(k^2+\lambda^2 s^2\right)v_k = 0$$

in the limit $s \to 1/2$, the Frobenius method of chapter 5 can be used. This was already done in section 9.2 for general dimensions n (see equation (9.2)). Substituting $n = 2$ yields the correct specialization for the case at hand. So, the first few terms in the local expansions of the solutions are

$$\begin{aligned} y_0(s) &:= 1 - \frac{8}{9}\left(4k^2+\lambda^2\right)\left(s-\frac{1}{2}\right)^3 + \frac{8}{5}(4k^2+\lambda^2)\left(s-\frac{1}{2}\right)^4 + O[5] \\ y_{3/2}(s) &:= \left(s-\frac{1}{2}\right)^{3/2} - \frac{9}{10}\left(s-\frac{1}{2}\right)^{5/2} + \frac{69}{56}\left(s-\frac{1}{2}\right)^{7/2} + O[9/2] \end{aligned} \tag{11.15}$$

From these two real solutions, a complex solution that satisfies the Wronskian normalization condition

$$W[v_k, v_k^*] = -2i\hbar\frac{\sqrt{4s^2-1}}{s} \tag{11.16}$$

can be constructed. Let $\tilde{v}_k$ be

$$v_k(s) := c_1 y_0(s) + i c_2 y_{3/2}(s). \tag{11.17}$$

Plugging this into $W[\tilde{v}_k, \tilde{v}_k^*]$ yields

$$ic_1c_2\left(-3\left(s-\frac{1}{2}\right)^{1/2} + \frac{9}{2}\left(s-\frac{1}{2}\right)^{3/2} - \frac{69}{8}\left(s-\frac{1}{2}\right)^{5/2}\right) + O[7/2].$$

On the other hand, expanding the right hand side of equation (11.16) to the same order yields

$$i\hbar\left(-8\left(s-\frac{1}{2}\right)^{1/2} + 12\left(s-\frac{1}{2}\right)^{3/2} - 23\left(s-\frac{1}{2}\right)^{5/2}\right) + O[7/2].$$

This implies that in order for $\tilde{v}_k$ to satisfy the Wronskian condition, the constants c_1 and c_2 must satisfy

$$c_1c_2 = \frac{8}{3}\hbar.$$

One possible choice is to let $c_1 = \sqrt{\hbar}$ and $c_2 = (8/3)\sqrt{\hbar}$. That is,

$$\tilde{v}_k = \sqrt{\hbar}\left(y_0 + i\frac{8}{3}y_{3/2}\right). \tag{11.18}$$

Now, while this is a solution to the mode equation, appropriately normalized, it is not necessarily the one that minimizes the energy

$$E_k(s) := \frac{1}{2}\frac{s^2}{4s^2-1}\left|\dot{v}_k\right|^2 + \frac{1}{2}\omega_k^2\left|v_k\right|^2 \tag{11.19}$$

in the limit of $s \to 1/2$. To find such the minimum energy mode function let

$$v_k := \alpha\tilde{v}_k + \beta\tilde{v}_k^*, \tag{11.20}$$

where α and β satisfy the Bogolyubov condition

$$|\alpha|^2 - |\beta|^2 = 1 \tag{11.21}$$

but are otherwise undetermined. This guarantees that v_k satisfies the Wronskian normalization condition (see section 10.4). Plugging (11.18) into (11.20) yields

$$v_k = \sqrt{\hbar}\left((\alpha+\beta)y_0 + i\frac{8}{3}(\alpha-\beta)y_{3/2}\right). \tag{11.22}$$

Then plugging this expression for v_k into the energy function (11.19), and taking the limit $s \to 1/2$ results in

$$E_k = \frac{\sqrt{\hbar}}{2}\left(|\alpha - \beta|^2 + \omega_k^2\big|_{s=1/2}\,|\alpha + \beta|^2\right). \tag{11.23}$$

Minimizing this expression subject to the constraint that $|\alpha^2| - |\beta|^2 = 1$ will yield the choice of mode function associated with the instantaneous minimum energy vacuum. However, before setting out to minimize said expression, it can be simplified a little. Since the energy expression (11.19) only depends on the modulus of v_k, any mode function that minimizes the expression can be multiplied by an overall phase factor without changing the value of E_k. This allows the freedom of letting one the two constants be real, say $\alpha = \alpha^*$. Writing $\beta = \beta_r + i\beta_i$ the problem can now be phrased as minimizing

$$\frac{\sqrt{\hbar}}{2}\left(\alpha^2 + \beta_r^2 + \beta_i^2 - 2\alpha\beta_r + \omega_k^2\big|_{s=1/2}\,(\alpha^2 + \beta_r^2 + \beta_i^2 + 2\alpha\beta_r)\right) \tag{11.24}$$

subject to the constraint

$$\alpha^2 - \beta_r^2 - \beta_i^2 - 1 = 0. \tag{11.25}$$

Using the typical method of Lagrange multipliers leads to the solution

$$\begin{aligned} \alpha &= \frac{1+\omega_k}{2\sqrt{\omega_k}} \\ \beta_r &= \frac{1-\omega_k}{2\sqrt{\omega_k}} \\ \beta_i &= 0, \end{aligned}$$

where for notational simplicity, the evaluation of ω_k at $s = 1/2$ is left implicit. So the mode function that minimizes the energy in the limit $s \to 1/2$ is

$$\begin{aligned} v_k &= \left(\frac{1+\omega_k}{2\sqrt{\omega_k}}\right)\tilde{v}_k + \left(\frac{1-\omega_k}{2\sqrt{\omega_k}}\right)\tilde{v}_k^* \\ &= \sqrt{\frac{\hbar}{\omega_k}}\left(y_0 + i\omega_k\frac{8}{3}y_{3/2}\right). \end{aligned} \tag{11.26}$$

Before moving on to the next limit, it is instructive to check that the F_k function (11.11), which could potentially spoil the ladder harmonic oscillator construction (see

discussion at the end of section 10.5) is indeed zero when evaluated on the minimum energy mode function. Plugging the expansions for y_0 and $y_{3/2}$ into (11.26) gives

$$\begin{aligned} v_k &= \sqrt{\frac{\hbar}{\omega_k}}\left(1 + i\omega_k \frac{8}{3}\left(s-\frac{1}{2}\right)^{3/2} - i\omega_k \frac{12}{5}\left(s-\frac{1}{2}\right)^{5/2} - \frac{32}{9}\omega_k^2\left(s-\frac{1}{2}\right)^3\right. \\ &\quad \left. + i\omega_k \frac{23}{8}\left(s-\frac{1}{2}\right)^{7/2} + \frac{32}{5}\omega_k^2\left(s-\frac{1}{2}\right)^4 + O\left[9/2\right]\right) \end{aligned}$$

where once again the fact that ω_k is being evaluated at $s = 1/2$ is left implicit. This in turn is used to calculate

$$\begin{aligned} (v_k)^2 &= \frac{\hbar}{\omega_k}\left(1 + i16\omega_k\left(s-\frac{1}{2}\right)^{3/2} - i\frac{24}{5}\omega_k\left(s-\frac{1}{2}\right)^{5/2} - \frac{128}{9}\omega_k^2\left(s-\frac{1}{2}\right)^3\right. \\ &\quad \left. + i\frac{46}{7}\omega_k\left(s-\frac{1}{2}\right)^{7/2} + \frac{128}{5}\omega_k^2\left(s-\frac{1}{2}\right)^4 + O\left[9/2\right]\right) \end{aligned} \tag{11.27}$$

and

$$\begin{aligned} (\dot{v}_k)^2 &= \frac{\hbar}{\omega_k}\left(-16\omega_k^2\left(s-\frac{1}{2}\right) + 48\omega_k^2\left(s-\frac{1}{2}\right)^2 - i\frac{256}{3}\omega_k^3\left(s-\frac{1}{2}\right)^{5/2}\right. \\ &\quad \left. - 128\omega_k^2\left(s-\frac{1}{2}\right)^3 + i\frac{1664}{5}\omega_k^3\left(s-\frac{1}{2}\right)^{7/2} + O\left[4\right]\right). \end{aligned} \tag{11.28}$$

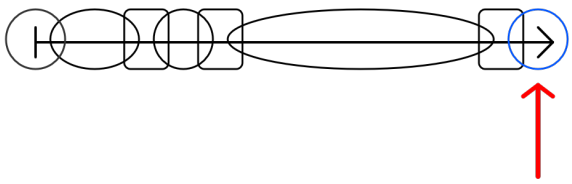
Plugging these expressions for $(v_k)^2$ and $(\dot{v}_k)^2$ into equation (11.11) and expanding the coefficient

$$\frac{s^2}{4s^2-1} \tag{11.29}$$

in terms of $(s - 1/2)$ yields precisely 0 up to order $(s - 1/2)^3$. This is as far as the expansions used for y_0 and $y_{3/2}$ are accurate. If higher order expansions were used, then the vanishing of F_k could be checked to higher orders.

$s \to \infty$ limit

Now the process must be repeated for the $s \to \infty$ limit. The first step is to obtain an asymptotic approximation to the solution to the mode equation in the appropriate limit. Using the methods from chapter 8 yields the following asymptotic series



solution

$$\begin{aligned} y_{\infty\pm} &= \exp\left(\pm i\lambda s^2 + \left(-\frac{1}{2} \pm i\left(\frac{k^2}{\lambda} - \frac{\lambda}{4}\right)\right)\ln(s)\right. \\ &\quad \left. + \left(-\frac{k^2}{4\lambda^2} \pm i\left(\frac{k^4}{8\lambda^3} + \frac{k^2}{16\lambda} + \frac{\lambda}{128} - \frac{3}{32\lambda}\right)\right)\frac{1}{s^2} + D(s)\right) \end{aligned} \tag{11.30}$$

where $D(s) = o(s^{-2})$ as $s \to \infty$. See appendix section D.3 for the details. This is already a pair of complex solutions, so let

$$\tilde{w}_k = c y_{\infty+}. \tag{11.31}$$

Plugging this $\tilde{w}_k$ into $W[\tilde{w}_k, \tilde{w}_k^*]$ yields

$$|c|^2\left(-i4\lambda s - i\left(\frac{k^2}{\lambda} - \frac{\lambda}{4}\right)\frac{2}{s} + o\left[s^{-2}\right]\right)\exp\left(-\ln(s) - \frac{k^2}{2\lambda^2}\frac{1}{s^2} + o\left[s^{-2}\right]\right)$$

which when expanded in powers of s, as $s \to \infty$ becomes

$$-i\lambda|c|^2\left(4 - \frac{1}{2s^2} + o[s^{-2}]\right).$$

On the other hand, expanding the right hand side of the Wronskian normalization condition (11.16) in terms of s yields

$$-i\hbar\left(4 - \frac{1}{2s^2} + o\left[s^{-2}\right]\right).$$

This implies that c satisfies

$$|c| = \sqrt{\frac{\hbar}{\lambda}}.$$

So a properly normalized mode function is

$$\tilde{w}_k = \sqrt{\frac{\hbar}{\lambda}} y_{\infty+}. \tag{11.32}$$

Again, this mode function can now be used to find the mode function that minimizes the Energy function. Letting

$$w_k = \alpha\tilde{\omega}_k + \beta\tilde{\omega}_k^* \tag{11.33}$$

where α and β satisfy $|\alpha|^2 - |\beta|^2 = 1$. Plugging the expression (11.33) into equation (11.19) order results in

$$E_k = \left(|\alpha|^2 + |\beta|^2\right)\left(\frac{1}{2}\frac{s^2}{4s^2-1}|\dot{\tilde{w}}_k|^2 + \frac{1}{2}\omega_k^2|\tilde{w}_k|^2\right) + 2\mathrm{Re}\left(\beta^*\alpha\left(\frac{1}{2}\frac{s^2}{4s^2-1}(\dot{\tilde{w}}_k)^2 + \frac{1}{2}\omega_k^2(\tilde{w}_k)^2\right)\right). \tag{11.34}$$

As before, α can be assumed to be real. So let $\alpha = \alpha^*$ and $\beta = \beta_r + i\beta_i$. Plugging in the definition of $\tilde{\omega_k}$, expanding in the limit $s \to \infty$, and keeping only the most dominant terms multiplying the undetermined coefficients yields

$$E_k = \left(\alpha^2 + |\beta|^2\right)\lambda^2 s - \frac{2\alpha}{s^3}\left(\beta_r(c_1\cos\theta + c_2\sin\theta) + \beta_i(c_1\sin\theta - c_2\cos\theta)\right)$$

where $\theta = 2\lambda s^2 + O[\ln(s)]$. The first terms grows unbounded in the limit $s \to \infty$, while the second term goes to zero, so minimizing this expression is tantamount to minimizing the first term. There the problem becomes finding the minimum of

$$\left(\alpha^2 + |\beta|^2\right) \tag{11.35}$$

subject to $\alpha^2 - |\beta|^2 = 1$. The solution is $\alpha = 1$ and $\beta = 0$. This means that the minimum energy vacuum near ∞ is that associated with the mode function

$$w_k = \sqrt{\frac{\hbar}{\lambda}}y_{\infty+}. \tag{11.36}$$

11.3 Matching mode functions

Having obtained an asymptotic approximation v_k associated with the instantaneous minimum energy vacuum in the limit $s \to 1/2$ and an asymptotic approximation w_k associated with the instantaneous minimum energy vacuum in the limit $\to \infty$, the last thing that needs to be done in order to answer the question of particle creation is to discover the relationship between these functions. That is, the next step is to obtain the coefficients α and β such that

$$w_k = \alpha v_k + \beta v_k^*. \tag{11.37}$$

So, the problem becomes that of matching the $s \to 1/2$ approximation

$$v_k = \sqrt{\frac{2\hbar}{\omega_k}} \left(y_0 + i\omega_k \frac{8}{3} y_{3/2} \right) \tag{11.38}$$

to the $s \to \infty$ approximation

$$w_k = \sqrt{\frac{\hbar}{\lambda}} y_{\infty+}. \tag{11.39}$$

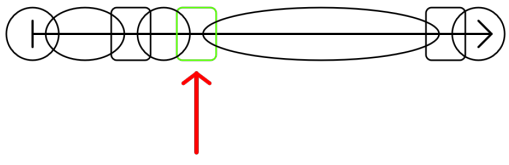
To this end a global geometric optics approximation is used, as was done in chapter 9. In other words, a global asymptotic approximation, in the limit $\epsilon := k^{-1}$ goes to zero, is matched to both the v_k approximation and the w_k approximation, thereby allowing a match between the v_k and w_k themselves. As before, it does not suffice to simply take the $s \to 1/2$ limit of the global solution, and the $\epsilon \to 0$ limit of the local local solution, and attempt to match the results. An intermediate approximation is needed that can bridge them together. Similarly, another intermediate solution is needed to match the global $\epsilon \to 0$ with the $s \to \infty$ solution.

The global $\epsilon \to 0$ approximation to the mode function was already obtained in section 9.3, albeit for the case of general dimension and in the Riemannian domain. It turns out that simply setting $n = 2$ and re-writing the expression assuming $s > 1/2$ instead of $s < 1/2$ yields the correct approximation for the case at hand. Namely,

$$S_{\epsilon\pm} = \exp\left(\pm \frac{i}{\epsilon} \left(\sqrt{4s^2 - 1} - \tan^1\left(\sqrt{4s^1 - 1} \right) \right) \pm i \frac{\epsilon\lambda^2}{24} \left(4s^2 - 1\right)^{3/2} - \frac{\epsilon^2\lambda^2}{16}(4s^2 - 1) + O[\epsilon^3] \right). \tag{11.40}$$

Matching $s \to 1/2$ with $\epsilon \to 0$ via Intermediate Approximation

As mentioned above, an intermediate approximation (see section 9.4) is needed to bridge the local $s \to 1/2$ and global $\epsilon \to 0$.



So, starting from the mode equation (with $\epsilon = k^{-1}$)

$$S'' - \frac{1}{s(4s^2 - 1)} S' + \left(\frac{1}{\epsilon^2} \frac{(4s^2 - 1)}{s^2} + \lambda^2(4s^2 - 1) \right) S = 0, \tag{11.41}$$

the first step is to perform the *stretching transformation*

$$s = \epsilon^\alpha r + \frac{1}{2} \tag{11.42}$$

and let $U(r(s)) = S(s)$. This yields

$$U'' - \frac{1}{2r(1+\epsilon^\alpha r)(1+2\epsilon^\alpha r)}U' + \left(16\epsilon^{3\alpha-2}\frac{r(1+\epsilon^\alpha r)}{(1+2\epsilon^\alpha r)^2} + 4\lambda^2\epsilon^{3\alpha}r(1+\epsilon^\alpha r)\right)U = 0. \tag{11.43}$$

The distinguished limit which takes into account the most dominant term in both the local $s \to 1/2$ dominant balance and the $s \to \infty$ dominant balance is $\alpha = 2/3$. In this case, the (stretched) mode equation becomes

$$U'' + PU' + QU = 0 \tag{11.44}$$

where

$$\begin{aligned} P &= -\frac{1}{2r(1+\epsilon^{2/3}r)(1+2\epsilon^{2/3}r)} \\ &\sim -\frac{1}{2r} + \frac{3}{2}\epsilon^{2/3} - \frac{7}{2}r\epsilon^{4/3} + \frac{15}{2}r^2\epsilon^2 + O\left[\epsilon^{8/3}\right] \qquad (11.45) \\ Q &= 16\frac{r(1+\epsilon^{2/3}r)}{(1+2\epsilon^{2/3}r)^2} + 4\lambda^2\epsilon^2 r(1+\epsilon^{2/3}r) \\ &\sim 16r - 48r^2\epsilon^{2/3} + 128r^3\epsilon^{4/3} + (4r\lambda^2 - 320r^4)\epsilon^2 + O\left[\epsilon^{8/3}\right]. \qquad (11.46) \end{aligned}$$

The three term dominant balance is

$$U'' - \frac{1}{2r}U' \sim -16rU \tag{11.47}$$

with solution

$$U_\pm = \exp\left(\pm i\frac{8}{3}r^{3/2} + B\right) \tag{11.48}$$

where $B = o(1)$ as $\epsilon \to 0$. Plugging the solution back into the equation, and keeping the next order terms in the expansions of P and Q yields

$$B'' + (B')^2 + \left(-\frac{1}{2r} \pm i8r^{1/2} + \frac{3}{2}\epsilon^{2/3}\right)B' \pm 6i\epsilon^{2/3}r^{1/2} - 48\epsilon^{2/3}r^2 = 0 \tag{11.49}$$

At this point, instead of solving the B dominant balance, which is very difficult, the work is split into two regimes: the $r \to 0$ and $r \to \infty$ limits. This leads to easier differential equations to solve and still allows the matching procedure to go through (see appendix C).

$r \to 0$ **regime**

In the limit $r \to 0$, the only consistent dominant balance is

$$B'' - \frac{1}{2r}B' \sim \mp 6i\epsilon^{2/3}r^{1/2}$$

whose solution is

$$\mp \frac{12}{5}i\epsilon^{2/3}r^{5/2}.$$

All together, the intermediate approximation in the $r \to 0$ regime is

$$U_\pm = \exp\left(\pm i\frac{8}{3}r^{3/2} \mp \frac{12}{5}i\epsilon^{2/3}r^{5/2} + B_2 + C_1\right)$$

where B_2 is proportional to $\epsilon^{2/3}$ and $o(r^{5/2})$ as $r \to 0$, and C_1 is $o(\epsilon^{2/3})$ as $\epsilon \to 0$. The solution is extended a bit further in appendix section (D.1), to obtain

$$U_\pm = \exp\left(\pm i\frac{8}{3}r^{3/2} \mp i\frac{12}{5}\epsilon^{2/3}r^{5/2} \pm \frac{23}{7}i\epsilon^{4/3}r^{7/2} + C_2 + D_1\right) \tag{11.50}$$

where C_2 is proportional to $\epsilon^{4/3}$ and $o(r^{7/2})$, and D is $o(\epsilon^{4/3})$.

$r \to \infty$ **regime**

In this regime there is a consistent two term dominant balance, namely

$$i8r^{1/2}B' \sim \pm 48\epsilon^{2/3}r^2,$$

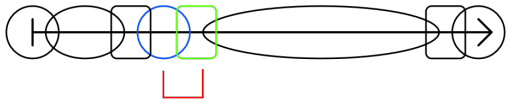
whose solution is

$$\mp \frac{12}{5}i\epsilon^{2/3}r^{5/2}.$$

Put together with the solution for the A term (11.48) results in the approximation

$$U_\pm = \exp\left(\pm i\frac{8}{3}r^{3/2} \mp \frac{12}{5}i\epsilon^{2/3}r^{5/2} + B_2 + C_1\right) \tag{11.51}$$

where B_2 is proportional to $\epsilon^{2/3}$ and $o(r^{5/2})$ as $r \to \infty$, and C_1 is $o(\epsilon^{2/3})$ as $\epsilon \to 0$. Notice that up to this order it agrees with the expansion in the regime $r \to 0$. In appendix section (D.1), the approximation is extended further. Even in the next order it continues to agree with the $r \to 0$ regime expansion. It is interesting that the same dominant terms arise for both the $r \to 0$ and $r \to \infty$ regime. This is not the case for $n > 2$ (see section 9.4).



Matching $r \to 0$ regime with $s \to 1/2$ Approximation

The next step is matching the approximate solution obtained in the limit $s \to \infty$ with the $r \to 0$ regime of the intermediate approximation. Performing the stretching transformation $s = \epsilon^{2/3} r + 1/2$ on the $s \to 1/2$ solutions $y_0(s)$ and $y_{3/2}(s)$ from equation (11.15), and then expanding about the limit $\epsilon \to 0$ yields

$$\begin{aligned} y_0(s(r)) &= 1 - \frac{32}{9} r^3 + \frac{32}{5} \epsilon^{2/3} r^4 + o[\epsilon^{2/3}] \\ y_{3/2}(s(r)) &= \epsilon r^{3/2} - \frac{9}{10} \epsilon^{5/3} r^{5/2} + \frac{69}{56} \epsilon^{7/3} r^{7/2} - \frac{32}{27} \epsilon r^{9/2} + o[\epsilon]. \end{aligned} \tag{11.52}$$

Plugging these into the mode function $y_0 + i\omega_k|_{1/2}(8/3)y_{3/2}$ and expanding the term $\omega_k|_{1/2}$ gives

$$1 + \frac{8}{3} i r^{3/2} - \frac{32}{9} r^3 - \frac{256}{81} i r^{9/2} - \frac{12}{5} i \epsilon^{2/3} r^{5/2} + \frac{32}{5} \epsilon^{2/3} r^4 + \frac{23}{7} i r^{7/2} \epsilon^{4/3} + o[\epsilon^{4/3}]$$

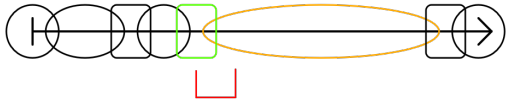
On the other end, taking the intermediate solution in the $r \to 0$ regime and expanding about $r = 0$ results in

$$\begin{aligned} S(s(r))_\pm &=: U_\pm(r) \\ &= 1 \pm \frac{8}{3} i r^{3/2} \mp \frac{12}{5} i \epsilon^{2/3} r^{5/2} - \frac{32}{9} r^3 \pm \frac{23}{7} i \epsilon^{4/3} r^{7/2} + \frac{32}{5} \epsilon^{2/3} r^4 + o[r^{7/2}] + o[\epsilon^{2/3}]. \end{aligned}$$

Comparing the two expansions leads to match

$$S_\pm = y_0 \pm \frac{8}{3} i \, \omega_k|_{s=1/2} \, y_{3/2}. \tag{11.53}$$

Match $r \to \infty$ regime with $s \to \infty$ Approximation



Now, the other regime of the intermediate approximation must be matched with the global $\epsilon \to 0$ approximation. Plugging in the stretching transformation into the global approximation (11.40) gives

$$\begin{aligned} S_{\epsilon\pm}(s(r)) &= \exp\Bigg(\pm \frac{i}{\epsilon}\left(2\epsilon^{1/3}r^{1/2}\sqrt{1+\epsilon^{2/3}r} - \tan^{-1}\left(2\epsilon^{1/3}r^{1/2}\sqrt{1+\epsilon^{2/3}r}\right)\right) \\ &\quad \pm i\frac{\epsilon^2\lambda^2}{3}r^{3/2}\left(1+\epsilon^{2/3}r\right)^{3/2} - \frac{\epsilon^2\lambda^2}{16}\left(4\epsilon^{4/3}r^2 + 4\epsilon^{2/3}r + 1\right) + ...\Bigg). \end{aligned}$$

As before, the $\epsilon \to 0$ must be taken again. However, since both approximations that are being matched right now are exponentials, it is easier to just work with the argument. So taking the $\epsilon \to 0$ limit of the argument of the above exponential yields

$$S_{\epsilon\pm} \sim \exp\left(\pm i\frac{8}{3}r^{3/2} \mp i\frac{12}{5}\epsilon^{2/3}r^{5/2} \pm \frac{23}{7}i\epsilon^{4/3}r^{7/2} + O[\epsilon^2]\right)$$

which, up to the order expanded, is precisely the intermediate solution in the $r \to \infty$ regime. This means that

$$S_{\epsilon\pm} = S_{\pm}. \tag{11.54}$$

Results

Having matched the intermediate approximation to both the local $s \to 1/2$ and the global $\epsilon \to 0$, the results can be combined to obtain the following:

$$v_k := \sqrt{\frac{\hbar}{\omega_k|_{s=1/2}}}\left(y_0 + i\,\omega_k|_{s=1/2}\,\frac{8}{3}y_{3/2}\right) \to \sqrt{\frac{\hbar}{\omega_k|_{s=1/2}}}\,(S_{\epsilon+}) \tag{11.55}$$

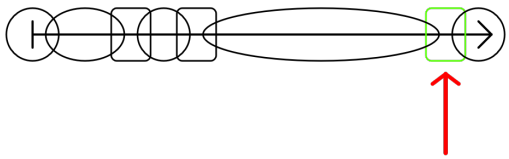
where in this case the symbol $\to$ is being used to convey the "matching" property. This concludes the first half of the matching problem.

Matching $\epsilon \to 0$ with $s \to \infty$ via Intermediate Approximation

Once more, to accomplish this match an intermediate approximation is required.

Starting from the same mode equation

$$S'' - \frac{1}{s(4s^2-1)}S' + \left(\frac{1}{\epsilon^2}\frac{(4s^2-1)}{s^2} + \lambda^2(4s^2-1)\right)S = 0,$$



this time the stretching transformation is

$$s = \frac{r}{\epsilon^\alpha}$$

for some $\alpha > 0$. Letting $U(r(s)) = S(s)$, the mode equation becomes

$$U'' - \frac{\epsilon^{2\alpha}}{r\left(4r^2 - \epsilon^{2\alpha}\right)}U' + \left(\frac{1}{\epsilon^{2(1+\alpha)}}\left(\frac{4r^2 - \epsilon^{2\alpha}}{r^2}\right) + \frac{\lambda^2}{\epsilon^{4\alpha}}\left(4r^2 - \epsilon^{2\alpha}\right)\right)U = 0$$

Here, the distinguished limit is $\alpha = 1$. Plugging this choice of α into the equation above yields

$$U'' - \frac{\epsilon^2}{r(4r^2 - \epsilon^2)}U' + \frac{1}{\epsilon^4}\left(4(1 + \lambda^2 r^2) - \frac{\epsilon^2(1 + \lambda^2 r^2)}{r}\right)U = 0.$$

As usual, it is convenient to let $U = \exp(A)$, which results in

$$A'' + (A')^2 - \frac{\epsilon^2}{r(4r^2 - \epsilon^2)}A' + \frac{1}{\epsilon^4}\left(4(1 + \lambda^2 r^2) - \frac{\epsilon^2(1 + \lambda^2 r^2)}{r}\right) = 0. \tag{11.56}$$

The consistent dominant balance is

$$(A')^2 \sim -\frac{1}{\epsilon^4}4(1 + \lambda^2 r^2)$$

whose solution is

$$A = \pm\frac{i}{\epsilon^2}\left(r\sqrt{1 + \lambda^2 r^2} - \frac{1}{\lambda}\ln\left(\sqrt{1 + \lambda^2 r^2} - \lambda r\right)\right) + B$$

where $B = o(\epsilon^{-2})$. Plugging this solution back into (11.56) yields

$$B'' + (B')^2 + \left(\pm\frac{4i}{\epsilon^2}\sqrt{1 + r^2\lambda^2} - \frac{\epsilon^2}{r(4r^2 - \epsilon^2)}\right)B'$$
$$+\frac{1}{\epsilon^2}\left(\pm\frac{2i\lambda^2 r}{\sqrt{1 + \lambda^2 r^2}} - \frac{1 + \lambda^2 r^2}{r^2}\right) \mp \frac{2i\sqrt{1 + \lambda^2 r^2}}{r(4r^2 - \epsilon^2)} = 0$$

At this point, the work could split into two regimes, $r \to 0$ and $r \to \infty$, as was done in the previous matching. However, in this case there is no need; the expansion

can be continued without resorting to taking said r limits. In particular, the dominant balance for B is

$$\pm \frac{4i}{\epsilon^2}\sqrt{1+\lambda^2 r^2}B' \sim -\frac{1}{\epsilon^2}\left(\pm\frac{2i\lambda^2 r}{\sqrt{1+\lambda^2 r^2}} - \frac{1+\lambda^2 r^2}{r^2}\right),$$

whose solution is

$$B = \pm\frac{i}{4}\left(\frac{\sqrt{1+\lambda^2 r^2}}{r} - \lambda \sinh^{-1}(\lambda r)\right) - \frac{1}{4}\ln(1+\lambda^2 r^2).$$

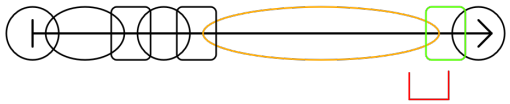
Putting the A and B solutions together yields

$$\begin{aligned} U_\pm &= \exp\left(\pm\frac{i}{\epsilon^2}\left(r\sqrt{1+\lambda^2 r^2} - \frac{1}{\lambda}\ln\left(\sqrt{1+\lambda^2 r^2} - \lambda r\right)\right)\right. \\ &\quad \left.\pm\frac{i}{4}\left(\frac{\sqrt{1+\lambda^2 r^2}}{r} - \lambda\sinh^{-1}(\lambda r)\right) - \frac{1}{4}\ln(1+\lambda^2 r^2) + C\right) \end{aligned} \tag{11.57}$$

where C is $o[1]$. Appendix (D.2) contains the derivation of the next term, which turns out to be

$$C = \pm\frac{i\epsilon^2}{192}\left(\frac{(1+\lambda^2 r^2)^{3/2}}{r^3} - \frac{4\lambda^2 r}{\sqrt{1+\lambda^2 r^2}} - \frac{20\lambda^2 r}{(1+\lambda^2 r^2)^{3/2}}\right).$$

Matching with $\epsilon \to 0$ approximation



Since all the relevant approximations being matched are exponentials, it is sufficient to work with the arguments. Starting with the global $\epsilon \to 0$ approximation, plugging the stretching transformation $s = r/\epsilon$ and expanding about $\epsilon \to 0$ again yields

$$\begin{aligned} Arg\left[S_{\epsilon\pm}(s(r))\right] &= \pm\frac{i}{\epsilon^2}\left(2r + \frac{\lambda^2 r^3}{3} + ...\right) \mp i\frac{\pi}{2\epsilon} \\ &\quad + \left(\pm\frac{i}{4r} \mp i\frac{r\lambda^2}{8} - \frac{\lambda^2 r^2}{4} + ...\right) \pm i\epsilon^2\left(\frac{1}{192r^3} + \frac{\lambda^2}{128r} + ...\right). \end{aligned}$$

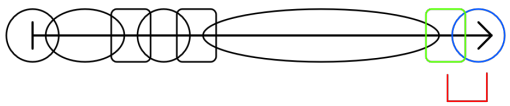
On the other hand, taking the intermediate approximation (11.57) and expanding about $r \to 0$ results in

$$\begin{aligned} Arg\left[U_{\pm}(r)\right] &= \pm\frac{i}{\epsilon^2}\left(2r + \frac{\lambda^2 r^3}{3} + ...\right) \\ &\quad + \left(\pm\frac{i}{4r} \mp i\frac{r\lambda^2}{8} - \frac{\lambda^2 r^2}{4} + ...\right) \pm i\epsilon^2\left(\frac{1}{192r^3} + \frac{\lambda^2}{128r} + ...\right) \\ &= Arg[S_{\epsilon\pm}(s(r))] \pm i\frac{\pi}{2\epsilon}. \end{aligned} \tag{11.58}$$

This in turn means that

$$U_{\pm}(r) = S_{\epsilon\pm}(s(r))\exp\left(\pm i\frac{\pi}{2\epsilon}\right) \tag{11.59}$$

Matching with $s \to \infty$ approximation



Now to match the other end, the stretching transformation is plugged into the local $s \to 0$ approximation (equation (11.30)), and the result is expanded about $\epsilon \to 0$. This becomes

$$\begin{aligned} Arg\left[y_{\infty\pm}(s(r))\right] &= \mp\frac{i\ln(\epsilon)}{\epsilon^2}\frac{1}{\lambda} \pm \frac{i}{\epsilon^2}\left(\lambda r^2 + \frac{1}{\lambda}\ln(r) + \left(\frac{1}{8\lambda^3}\right)\frac{1}{r^2} + ...\right) \\ &\quad + \left(\frac{1}{2} \pm \frac{i\lambda}{4}\right)\ln(\epsilon) + \left(-\left(\frac{1}{2} \pm \frac{i\lambda}{4}\right)\ln(r) - \left(\frac{1}{4\lambda^2} \mp \frac{i}{16\lambda}\right)\frac{1}{r^2} + ...\right) \\ &\quad \pm i\epsilon^2\left(\left(-\frac{3}{32\lambda} + \frac{\lambda}{128}\right)\frac{1}{r^2} + ...\right) + o(\epsilon^2) \end{aligned}$$

Then, taking expanding the intermediate solution about $r \to \infty$ yields

$$\begin{aligned} Arg[U_{\pm}(r)] &= \pm\frac{i}{\epsilon^2}\left(\lambda r^2 + \frac{1}{\lambda}\ln(r) + \left(\frac{1}{2\lambda} + \frac{\ln(2\lambda)}{\lambda}\right) + \frac{1}{8\lambda^3}\frac{1}{r^2} + ...\right) \\ &\quad + \left(-\left(\frac{1}{2} \pm \frac{\lambda}{4}\right)\ln(r) - \frac{1}{2}\ln(\lambda) \pm \frac{i\lambda}{4}\left(1 - \ln(2\lambda)\right) - \left(\frac{1}{4\lambda^2} \mp \frac{i}{16\lambda}\right)\frac{1}{r^2} + ...\right) \\ &\quad \pm i\epsilon^2\left(-\frac{\lambda}{48} + \frac{\lambda^3}{192} + \left(-\frac{3}{32\lambda} + \frac{\lambda}{128}\right)\frac{1}{r^2} + ...\right). \end{aligned}$$

This implies that

$$\begin{aligned} Arg[U_{\pm}(r)] &= Arg\left[S_{\infty\pm}(s(r))\right] \pm \frac{i\ln(\epsilon)}{\epsilon^2}\frac{1}{\lambda} \pm \frac{i}{\epsilon^2}\left(\frac{1}{2\lambda} + \frac{\ln(2\lambda)}{\lambda}\right) \\ &\quad - \left(\frac{1}{2} \pm \frac{i\lambda}{4}\right)\ln(\epsilon\lambda) \pm \frac{i\lambda}{4}\left(1 - \ln(2)\right) \pm i\epsilon^2\left(-\frac{\lambda}{48} + \frac{\lambda^3}{192}\right) + O\left[\epsilon^3\right] \\ &= Arg\left[S_{\infty\pm}(s(r))\right] - \frac{1}{2}\ln(\lambda\epsilon) \pm i\Theta(\lambda,\epsilon) + O[\epsilon^3] \end{aligned}$$

where

$$\Theta(\lambda,\epsilon) := \frac{\ln(\epsilon)}{\epsilon^2}\frac{1}{\lambda} + \frac{1}{\epsilon^2}\left(\frac{1}{2\lambda} + \frac{\ln(2\lambda)}{\lambda}\right) - \frac{\lambda}{4}\left(\ln(\epsilon\lambda) + \ln(2) - 1\right) + \epsilon^2\left(-\frac{\lambda}{48} + \frac{\lambda^3}{192}\right).$$

In particular

$$U_{\pm}(r) = \frac{1}{\sqrt{\lambda\epsilon}}\exp\left(\pm i\Theta(\lambda,\epsilon)\right) y_{\infty\pm}(s(r))$$

Results

Now, putting these two intermediate matches together yields a match between the global $\epsilon \to 0$ approximation and the local $s \to \infty$ approximation, namely

$$S_{\epsilon\pm}(s) = \frac{1}{\sqrt{\lambda\epsilon}} e^{\pm i\bar{\Theta}(\lambda,\epsilon)} y_{\infty\pm}(s) \tag{11.60}$$

where $\bar{\Theta} = \Theta - \frac{\pi}{2\epsilon}$. This means in turn implies that

$$w_k \to \sqrt{\epsilon\hbar} e^{-i\bar{\Theta}(\lambda,\epsilon)} S_{\epsilon+} \tag{11.61}$$

11.4 Complete Matching Results

The mode function whose minimum excitation state minimizes the energy as $s \to 1/2$ is

$$v_k = \sqrt{\frac{\hbar}{\omega_k}}\left(y_0 + i\,\omega_k|_{1/2}\,\frac{8}{3}y_{3/2}\right).$$

where $\omega_k^2 = (k^2 + \lambda^2 s^2)$, and y_0 and $y_{3/2}$ are detailed in equation (11.15). Matching the $s \to 1/2$ approximation with the global $\epsilon \to 0$ approximation via an intermediate approximation resulted in

$$v_k := \sqrt{\frac{\hbar}{\omega_k|_{s=1/2}}}\left(y_0 + i\,\omega_k|_{s=1/2}\,\frac{8}{3}y_{3/2}\right) \to \sqrt{\frac{\hbar}{\omega_k|_{s=1/2}}}\left(S_{\epsilon+}\right)$$

At the other end of the s domain, the mode function associated with the minimum energy in the limit $s \to \infty$ is

$$w_k = \sqrt{\frac{\hbar}{\lambda}} y_{\infty +}.$$

where $y_{\infty +}$ is shown in equation (11.30). Matching the $\epsilon \to 0$ approximation to the $s \to \infty$ approximation via another intermediate approximation gave rise to

$$w_k \to \sqrt{\epsilon \hbar} e^{-i\bar{\Theta}(\lambda,\epsilon)} S_{\epsilon +}.$$

Putting these two matches together leads to the ultimately sought after match

$$\boxed{w_k \to \sqrt{\epsilon\, \omega_k|_{1/2}} e^{-i\bar{\Theta}} v_k.} \tag{11.62}$$

There are a couple of important things to note about this match.

- The matching relationship is not a strict equality. Instead it is only accurate to the highest order ϵ consistently used in the expansions. So the constant in front of v_k is only accurate up to order $\epsilon^2 = k^{-2}$.

- Furthermore, it is not the case that w_k is proportional to v_k, as equation (11.62) would have it seem. Instead, there is also a term proportional to v_k^* that cannot be obtained through the matching procedure performed. This happens because the matching was performed via the use of asymptotic expansions, and at every order only the most dominant terms are considered. Any term proportional to v_k will be dominant over any term proportional to v_k^*, in the $\epsilon \to 0$ limit. This is due to a combination of two facts. First, v_k matches $S_{\epsilon +}$ and v_k^* matches $S_{\epsilon -}$ (see equations (11.53) and (11.54)). Second, in the limit $\epsilon \to 0$, the global solution $S_{\epsilon -}$ is subdominant to the solution $S_{\epsilon +}$. This can be seen by comparing the leading terms in the $S_{\epsilon \pm}$ solutions (see equation (11.40)).

Given these two points, a more accurate representation of the match is

$$w_k = \left(1 + \frac{\lambda^2}{16k^2} + O[k^{-4}]\right) e^{-i\bar{\Theta}(\lambda,k)} v_k + \beta v_k^*$$

where β is an as yet undetermined complex constant that depends on k and λ. As described in section 10.4, the information about the particle content is contained in

the β constant that the matching procedure cannot capture. However, since both w_k and v_k were appropriately normalized according to the Wronskian condition, the transformation relating them are a Bogolyubov transformation thus they satisfy the condition

$$|\alpha|^2 - |\beta|^2 = 1.$$

This implies that

$$\boxed{|\beta|^2 = \frac{\lambda^2}{8k^2} + O[k^{-4}].} \tag{11.63}$$

So, as detailed in equation 10.43, this is the number of particles produced, per momentum mode k, by the Klein-Gordon field's interaction with gravity. The number of particles decreases as the square of the wavenumber. The immediate conclusion is that there *is* particle production in this cosmology. The method gives the high-frequency (high-energy) limit of the production spectrum; the production decreases proportional to $1/k^2$. [It is not surprising that the particle production vanishes as wavenumber $k \to \infty$. Cosmological particle production occurs when the dynamical metric interacts with the evolution of the wave. But when $k \approx \infty$, metric changes become infinitely slow during any one wave cycle, hence no particle production.] Since the analysis uses asymptotic expansions in the high k limit, I cannot extend the prediction to larger wavelength.

As far as the λ proportionality, it makes sense that if λ vanishes, the particle production should vanish as well. The constant λ represents the mass term, and this is what couples the field to gravity. To understand this relationship further, recall that in the non-dimensionalization step, s was defined as

$$s := \frac{t}{(f/c)}$$

so f/c is the relevant time scale of the problem. In a sense, it controls how much "time" is spent in the Riemannian domain. That is, in terms of the t coordinate, the universe transitions from Riemannian to Lorentzian when $t = (1/2)(f/c)$. Furthermore, λ is defined as

$$\lambda := \frac{\mu c f}{\hbar}$$

which can be rewritten as

$$\lambda = \frac{(\mu c^2)(f/c)}{\hbar}.$$

So it effectively compares the product of the "rest energy" in the field, μc^2 and the natural time scale of the problem to Planck's constant $\hbar$. It appears that, for a fixed mass, the more (Euclidean) time the universe spends brewing in a classically forbidden state, the greater the consequent particle production.

11.5 Aside: Back-reaction

Having computed the expected number of particles produced through the interaction of the Klein-Gordon field with the gravitational background, it is natural to consider the effect of these particles on the background itself. This is the typical back-reaction problem that occurs in many perturbation/semi-classical calculations. When working in the context of cosmology, the effect of the back-reaction is thought to be of major importance. These quantum field theoretical effects might help avoid cosmological singularities or contribute to the accelerated expansion of the universe.

The basic idea is to continue to work with a classical spacetime and a quantum field, but attempt to link their dynamical evolutions via a generalization of Einstein's equations

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{8\pi G}{c^4} T_{\mu\nu}.$$

The natural thing to try is to use the classical expression for $T_{\mu\nu}$ in terms of the Klein-Gordon field, and take its vacuum expectation value, then use that as the right hand side of Einstein's equations. Namely

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{8\pi G}{c^4} \langle T_{\mu\nu} \rangle. \tag{11.64}$$

This formal expression for $\langle T_{\mu\nu} \rangle$ can be shown to diverge, as expected, since it contains products of operator valued distributions on spacetime. Not only is this a problem for obtaining insight regarding the back-reaction problem, but at first glance it calls into question the validity of the QFTCS framework. For interesting discussions on the consistency of this back-reaction problem see Wald [1977] and Padmanabhan [1989].

This problem is also hinted at at by the result obtained in the previous section. If the fact that particle count goes like k^{-2} holds for higher dimensions, then, in the $3+1$ case, the total number of particles would like

$$\int \frac{1}{k^2} k^2 dk = \int dk. \tag{11.65}$$

This combined with the fact that the energy in k mode goes like $(k^2 + \lambda^2 s^2)^{1/2}$ leads to a potentially very serious back-reaction problem.

This particular divergence is not a new phenomenon. The infinite vacuum energy is present even in the case of quantum fields in Minkowski spacetime. However, in the case of non-gravitating physics, the argument typically goes as follows: The total value of the energy is not relevant, only differences in energy are relevant, so removing a background energy value (even if it is infinite) does not impact the physics. This is then done via the imposition of normal ordering - a very simple renormalization technique.

Lamentably, in QFTCS, the absence of a uniquely defined vacuum state, and its accompanying ladder operators make this renormalization via normal ordering not well defined. So to accomplish the same effect, different renormalization techniques are needed. Moreover, the more fundamental problem is that for a theory involving gravity, the total energy does matter, it affects the curvature of spacetime. So any renomarlization technique is somewhat suspect.

However, there are different approaches to successfully remove infinities from the expected value of the stress energy tensor, leaving behind a residue that might be a suitable candidate for the source term in the RHS of Einstein's equations. Point splitting techniques well known from regular quantum field theory can be used, as well as removing infinities via subtracting adiabatic solutions to the mode equation (using a WKB approximation much like the one used in the previous sections). Another approach is to use a DeWitt-Schwinger inverse mass expansion to compute an effective action (see Birrell and Davies [1984]). In some cases, these different techniques yield the same results, which lends confidence to the validity of the approximation.

There are several results that show that under suitable renormalization techniques, the back-reaction is small enough to render the QFTCS approximation consistent in certain situations, like particular regimes in a slow roll inflation models (see Kaya and Tarman [2012]), or a non-conformal scalar field in a de Sitter spacetime (see Pérez-Nadal et al. [2008]). In some cases, the contribution of the back-reaction can be suitably incorporated into to the overall dynamical evolution, like in certain FLRW spacetimes (see Koivisto and Prokopec [2011]).

Appendices

Appendix A: Matching Geodesics

To analyze geodesics, the usual place to start is with the affinely parameterized geodesic equation. For $\gamma : \mathbb{R} \to M$, the geodesics equation is typically expressed as

$$\frac{d^2\gamma^\rho}{d\lambda^2} + \Gamma^\rho{}_{\mu\nu}\frac{d\gamma^\mu}{d\lambda}\frac{d\gamma^\nu}{d\lambda} = 0$$

For the sake of simplicity in notation (and agreement with literature) let $\gamma^0(\lambda)$ be denoted $t(\lambda)$ and similarly $\gamma^1(\lambda)$ denoted $\theta(\lambda)$. Then the geodesic equation becomes

$$t'' + \frac{4t}{4t^2-1}\left(t'\right)^2 + \frac{t}{4t^2-1}\left(\theta'\right)^2 = 0 \tag{A.1}$$

$$\theta'' + \frac{2}{t}t'\theta' = 0 \tag{A.2}$$

The second equation can be integrated immediately to obtain

$$\theta' = \frac{\alpha}{t^2}$$

where α is an arbitrary (including potentially zero) constant on integration which encodes information about the spatial velocity.

It is often simplest to first consider just the null geodesics since the added null-like condition is a first order differential equation. For this toy model, that condition is

$$-(4t^2-1)\left(\theta'\right)^2 + t^2\left(t'\right) = 0$$

which can be rewritten as

$$\left(\theta'\right)^2 = \frac{(4t^2-1)}{t^2}\left(t'\right)^2.$$

The first thing to notice, which is completely expected, is that for $t < 1/2$ this equation cannot be satisfied since both t and θ are real functions of a real variable and the LHS is positive, while the RHS is negative. The same is true for time-like geodesics, where the equality would be replaced by a less-than symbol. This is simply a restatement of the fact that the metric g is Riemannian for $t < 1/2$ (positive-definite) and there is no light cone structure. So from the outset it is clear that attempting to understand the behavior of particles by solving the geodesic equation in both domains D_R and D_L and matching at the boundary is murky at best. Every

geodesic on the Lorentzian domain will have to be matched to a geodesic devoid of causal meaning in the Riemannian domain. However, analyzing geodesics a little further will still result in some fruitful insight.

Therefore, plugging the null condition into the integral of the second geodesic equation gives[1]

$$t't\sqrt{4t^2-1} = \alpha$$

which can be integrated to obtain

$$\frac{1}{12}\left(4t^2-1\right)^{3/2} = \alpha\lambda + \beta$$

where β is another arbitrary constant of integration. This solution can then be plugged back into the θ' equation to solve for θ in terms of λ. Imposing the arbitrary, but legal choice that the transition boundary ($t = 1/2$) is crossed when $\lambda = 0$, at $\theta = 0$, leads to the solution

$$\begin{aligned} t(\lambda) &= \frac{1}{2}\sqrt{(\alpha\lambda)^{2/3}+1} \\ \theta(\lambda) &= (\alpha\lambda)^{1/3} - \tan^{-1}\left((\alpha\lambda)^{1/3}\right) \end{aligned}$$

where α has been redefined to absorb a factor of 12. Figure A.1 below portrays the geodesic, with α set to 1 and λ ranging from 0 to 0.5.

The null geodesic behaves much like a Minkowski null geodesic does away from the transition. In the limit of λ approaches infinity, θ is approximately $2t$. On the other hand, close to the transition the null geodesics appears to be almost vertical. That is, there is almost no change in θ as t changes. The coordinate velocity $d\theta/dt$ in terms of the parameter λ is

$$\frac{d\theta}{dt} = \frac{\frac{d\theta}{d\lambda}}{\frac{dt}{d\lambda}} = \frac{2(\alpha\lambda)^{1/3}}{\sqrt{(\alpha\lambda)^{2/3}+1}}.$$

In terms of t, the coordinate velocity becomes

$$\frac{d\theta}{dt} = \frac{(4t^2-1)^{1/2}}{t}.$$

This geodesic cannot be extended to t less than $1/2$, so the singularity at $t = 0$ is of no particular concern. However, as t approaches $1/2$ from above, the coordinate

[1]Setting α to be zero implies $t' = \theta' = 0$. This is discarded as not being an interesting solution.

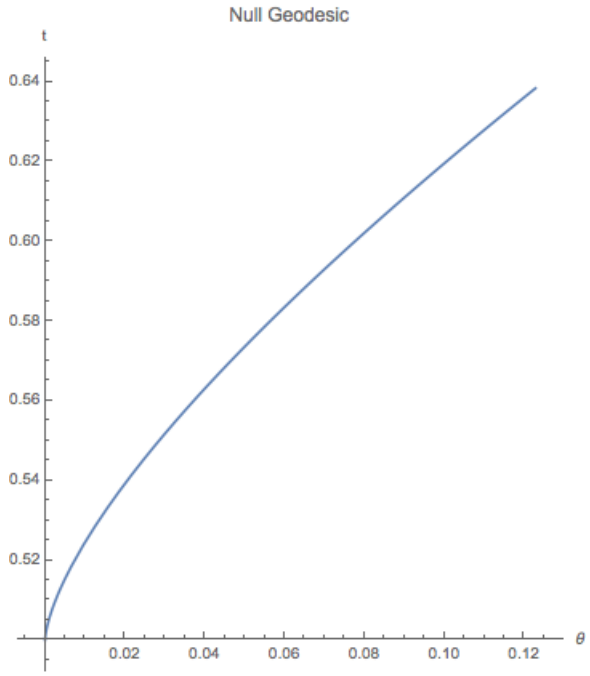


Figure A.1: Null geodesic for 2D toy model. Parameter λ takes values between 0 and 1/2, and constant α chosen to be 1.

velocity approaches 0. In other words the light cones squeeze shut. This implies that massive particles, on time-like geodesics are bunched closer and closer together (in terms of allowed speeds) the closer they are to the transition surface at $t = 1/2$. Or alternatively stated, all physical geodesics become null-geodesics.

In a sense, this is entirely an artifact of the choice of time coordinate. In terms of cosmic time the coordinate velocity is

$$\frac{d\theta}{d\tau} = \frac{d\theta}{dt}\frac{dt}{d\tau} = \frac{1}{t(\tau)} \tag{A.3}$$

- Nothing particularly bad happens at $t = 1/2$
- On the other hand, $t = 1/2$ is a boundary point for τ. Any statement involving τ at the point $t = 1/2$ is suspect.

Now, for general geodesics, still affinely parameterized, the only relevant equations

are

$$t'' + \frac{4t}{4t^2-1}\left(t'\right)^2 + \frac{t}{4t^2-1}\left(\theta'\right)^2 = 0$$
$$\theta'' + \frac{2}{t}t'\theta' = 0.$$

Once again, the second equation can be integrated $\theta' = \frac{\alpha}{t^2}$. Plugging this expression for θ' back in to the first equation results in an ordinary differential equation for $t(\lambda)$, in particular

$$t'' + \frac{4t}{4t^2-1}\left(t'\right)^2 + \frac{\alpha^2}{t^3(4t^2-1)} = 0.$$

By multiplying by $\sqrt{\pm\left(4t^2-1\right)}$ - where the sign is chosen depending on the domain under consideration - the equation can be rewritten as

$$\frac{d}{d\lambda}\left(\sqrt{\pm(4t^2-1)}\frac{dt}{d\lambda}\right) + \frac{\alpha^2}{t^3\sqrt{\pm(4t^2-1)}} = 0.$$

This equation cannot be solved in a closed form for arbitrary α. However, setting $\alpha = 0$, which means that $\theta(\lambda)$ is constant, makes it to where the equation for $t(\lambda)$ can be integrated. The result depends on the domain under consideration. In the Riemannian domain D_R the integrated equation becomes

$$\frac{1}{2}t\sqrt{-(4t^2-1)} + \frac{1}{4}\sin^{-1}(2t) = \beta\lambda + \zeta_R,$$

while in the Lorentzian domain D_L it becomes

$$\frac{1}{2}t\sqrt{4t^2-1} - \frac{1}{4}\ln\left(2t + \sqrt{4t^2-1}\right) = \beta\lambda + \zeta_L.$$

where β, ζ_R and ζ_L are arbitrary constants of integration. For continuity of $\lambda(t)$ at $t = 1/2$, the constants ζ_R and ζ_L are set to $\pi/8$ and 0 respectively. This leads to the following equation for λ in terms of t

$$\lambda = \begin{cases} \frac{1}{\beta}\left(\frac{1}{2}t\sqrt{-(4t^2-1)} + \frac{1}{4}\sin^{-1}(2t) - \frac{\pi}{8}\right) & 0 < t \leq \frac{1}{2} \\ \frac{1}{\beta}\left(\frac{1}{2}t\sqrt{4t^2-1} - \frac{1}{4}\ln\left(2t + \sqrt{4t^2-1}\right)\right) & \frac{1}{2} \geq t < \infty \end{cases}.$$

The following is a plot of $\lambda(t)$.

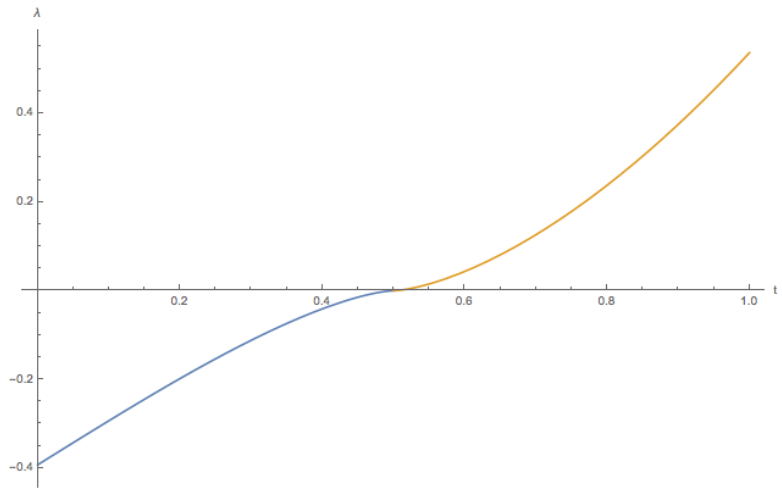


Figure A.2: Affine parameter λ as a function of coordinate t coordinate for a time-like geodesic. Parameter λ takes values between -1 and 1 as t goes from 0 to $1/2$, and constant β chosen to be 1.

Lamentably, the above expression for $\lambda(t)$ cannot be solved for t in term of λ, in either domain. In fact, notice that

$$\left.\frac{d\lambda}{dt}\right|_{1/2} = 0$$

regardless of whether $t = 1/2$ is approached from above or below. Therefore the inverse function theorem can't be used to guarantee the existence of a local inverse. However figure A.2 seems to indicate that the inverse function theorem fails in the manner $f(x) = x^3$ does, around $x = 0$, as opposed to the way $f(x) = x^2$ does. Perhaps there is still a well defined local inverse.

To further investigate the transition, I narrow in to the transition region, and focus on the part of the geodesic that is very close to $t = 1/2$. In other words, let

$$t = \frac{1}{2} + \epsilon(\lambda)$$

where $|\epsilon| << 1$. While this can be simply plugged into the expression for λ in terms of t, expanded to lowest order in ϵ and then solved for ϵ in terms of λ, it will be somewhat more useful to go back to the equation

$$t'' + \frac{4t}{4t^2 - 1}(t')^2 + \frac{\alpha^2}{t^3(4t^2 - 1)} = 0$$

and plug in the approximation there. This will reveal information about generic time-like geodesics and not only those with zero coordinate velocity. Plugging the approximation for t leads to

$$\epsilon'' + \frac{1+2\epsilon}{2\epsilon(\epsilon+1)}(\epsilon')^2 + \frac{\alpha^2}{\epsilon(\epsilon+1)\left(\frac{1}{2}+3\epsilon+6\epsilon^2+4\epsilon^3\right)} = 0.$$

Dropping all but the lowest order ϵ terms in each of the three summands leads to

$$\epsilon'' + \frac{1}{2\epsilon}(\epsilon')^2 + \frac{2\alpha^2}{\epsilon} = 0.$$

Even this simplified, local equation cannot be solved analytically. However, there are two conclusions that can be drawn form this equation without resorting to careful applications of asymptotic analysis. First, the case $\alpha = 0$ (zero coordinate velocity) reduces to

$$\epsilon'' + \frac{(\epsilon')^2}{2\epsilon} = 0.$$

Ignoring the ϵ equals constant solution and dividing the equation by ϵ' leads to

$$\frac{\epsilon''}{\epsilon'} + \frac{1}{2}\frac{\epsilon'}{\epsilon} = 0.$$

This equation can be immediately integrated, however the results depend on the sign ϵ and its derivatives. This can be summarized as

$$\ln(|\epsilon'|) + \frac{1}{2}\ln(|\epsilon|) = \zeta_1.$$

This already portends ill behavior at the transition since the solution will be a piece-wise defined function. In particular, the solution to this equation is

$$\epsilon = \begin{cases} -\left(\zeta_1\lambda + \zeta_2\right)^{2/3} & \epsilon < 0 \\ \left(\zeta_3\lambda + \zeta_4\right)^{2/3} & \epsilon > 0 \end{cases}$$

where the ζ_i for $i = 1..4$ are constants of integration, ζ_2 and ζ_4 being arbitrary and ζ_1 and ζ_3 necessarily non-zero. Without loss of generality ζ_2 and ζ_4 can be set to zero, thus making the transition happen at $\lambda = 0$. This is also enough to ensure continuity across the transition. Notice however that there is no (non-zero) choice of ζ_1 and ζ_3 that can make this solution even first differentiable. This is precisely as surmised from the exact t and α equal zero analysis earlier. In fact, setting

$$\zeta_1 = -\frac{3}{4}\beta, \qquad \zeta_3 = \frac{3}{4}\beta$$

the local solution takes the form

$$\epsilon = \begin{cases} -\left(-\frac{3}{4}\beta\lambda\right)^{2/3} & \lambda < 0 \\ \left(\frac{3}{4}\beta\lambda\right)^{2/3} & \lambda > 0 \end{cases}$$

which is precisely the local behavior of the exact (albeit not-invertible) solution obtained earlier, for λ in terms of t. This blowing up of the first derivative is troublesome since it is definitely not the behavior desired for a parameterized curve.

Now, when α is not set to zero, there isn't much progress to be made. However, rewriting the equation governing the evolution of ϵ as

$$\epsilon\epsilon'' + \frac{1}{2}(\epsilon')^2 + 2\alpha^2 = 0$$

does lead to some insight. In particular, since α is finite and positive, the term $\epsilon\epsilon''$ must remain finite and negative even as ϵ approaches zero, in order for the equation to be satisfied. The second term, which is proportional to $(\epsilon')^2$ cannot cancel out the third term, since it is also positive. This means that ϵ'' must go from negative infinity to positive infinity at transition. The second derivative does not exist when $\epsilon = 0$ hence the differential equation cannot be satisfied at that point. So a nonzero α cannot save the geodesic.

A.1 Non-Affinely Parametrized (NAP) Geodesics

From the analysis in the previous section, it is clear that the geodesics all break down at the transition surface. The solutions to the affineley parameterized geodesics are not even properly parameterized curves. This begs the question of whether or not the culprit is the affine parameter itself and not the geodesics. To answer this I turn to the non-affinely parameterized (NAP) geodesic equation.

$$\frac{d^2\gamma^\rho}{d\lambda^2} + \frac{1}{2}g^{\rho\sigma}\left(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu}\right)\frac{d\gamma^\mu}{d\lambda}\frac{d\gamma^\nu}{d\lambda} = h(\lambda)\frac{d\gamma^\rho}{d\lambda} \tag{A.4}$$

which written in the notation employed in the previous section becomes

$$\begin{aligned} t'' + \frac{4t}{4t^2-1}\left(t'\right)^2 + \frac{t}{4t^2-1}\left(\theta'\right)^2 &= ht' \\ \theta'' + \frac{2}{t}t'\theta' &= h\theta'. \end{aligned} \tag{A.5}$$

In the equations above, $h(\lambda)$ is an arbitrary function of λ that controls the parameterization of the geodesic. In particular $h = 0$ corresponds to an affine parameterization. Perhaps there are other choices of h, leading to alternative parameterizations, which result in geodesics that are well behaved at the transition.

As before, the second equation can be readily integrated, leading to

$$\theta' = \frac{\alpha}{t^2} e^{\int d\lambda h}$$

where α is an arbitrary constant of integration. Plugging this back into the first equation yields

$$t'' + \frac{4t}{4t^2 - 1}\left(t'\right)^2 + \frac{\alpha^2}{t^3(4t^2 - 1)} e^{2\int d\lambda h} = ht'.$$

Now, jumping straight into the local analysis, close to $t = 1/2$ by once again letting $t = 1/2 + \epsilon$ and considering only $|\epsilon| << 1$ leads to

$$\epsilon'' + \frac{1}{2\epsilon}(\epsilon')^2 + \frac{2\alpha^2}{\epsilon} e^{2\int d\lambda h} = h\epsilon'.$$

At this point I choose a form for the arbitrary function h. In particular I let

$$h(\lambda) = \frac{2}{\lambda}.$$

Further restricting the analysis to the zero coordinate velocity case ($\alpha = 0$) results in the following equation:

$$\epsilon'' + \frac{1}{2\epsilon}(\epsilon')^2 = \frac{2}{\lambda}\epsilon'.$$

This equation can be solved in the same way the affinely parameterized counterpart above was solved. The solution to this equation is

$$\epsilon = \begin{cases} -\left(\zeta_1\lambda^3 + \zeta_2\right)^{2/3} & \epsilon < 0 \\ \left(\zeta_3\lambda^3 + \zeta_4\right)^{2/3} & \epsilon > 0 \end{cases}$$

where again ζ_2 and ζ_4 are arbitrary constants of integration, while ζ_1 and ζ_3 are necessarily non-zero. Setting ζ_2 and ζ_4 to zero makes the transition happen at $\lambda = 0$, which ensures continuity across the transition. After redefining some arbitrary constants of integration, the solution becomes

$$\epsilon = \begin{cases} -\zeta_L\lambda^2 & \epsilon < 0 \\ \zeta_R\lambda^2 & \epsilon > 0 \end{cases}.$$

This time, this solution has a continuous first derivative. The second derivative, on the other hand, cannot be made continuous by choosing the constants of integration appropriately. However, the second derivative is finite (with a finite jump). This has the added benefit that whereas in the affinely parameterzied case the equation for ϵ was divergent at $\epsilon = 0$, now there is a sense in which it is valid and moreover satisfied at that point. The equation is valid in the limit that ϵ approaches 0. Granted that limit depends on whether the transition is approached from above or from below.

Furthermore, the discontinuity in the second derivative can be pushed to higher and higher derivatives via judicious reparameterizations - or equivalently different choices of the arbitrary function h in the NAP geodesic equation. Although it is unclear, and perhaps unlikely that the discontinuity can be completely removed.

Appendix B: Hamiltonian Approach to Geodesics

The characterization of geodesics - which are the paths freely-falling particles follow - in terms of the extremization of the length functional is effectively a Lagrangian formulation of particle mechanics. The salient difference is that the path is parameterized by an arbitrary parameter λ as opposed to some universal notion of time, which instead is associated with a degree of freedom $\gamma^0(\lambda)$ (i.e. the time component of the coordinate expression of the path γ). There is a natural Hamiltonian counterpart to this Lagrangian formulation. This is a key element of the discussion of the correspondence between waves and particles (see chapter 3).

The length functional is the action

$$S[\gamma^\mu] = \int_{\lambda_0}^{\lambda_1} d\lambda \, L\left(\gamma^\nu, \frac{d\gamma^\nu}{d\lambda}\right),$$

where the Lagrangian is

$$L\left(\gamma^\nu, \frac{d\gamma^\nu}{d\lambda}\right) = \sqrt{(sgn) g_{\mu\nu}(\gamma^\alpha) \frac{d\gamma^\mu}{d\lambda} \frac{d\gamma^\nu}{d\lambda}}.$$

The $g_{\mu\nu}$ are just given functions of the generalized coordinates γ^α. Proceeding in the usual fashion, from a Lagrangian formulation to a Hamiltonian formulation, the first step is to obtain the canonical momenta conjugate to the generalized coordinates γ^μ. In particular these are

$$\begin{aligned} p_\alpha &:= \frac{\partial L}{\partial\left(\frac{d\gamma^\alpha}{d\lambda}\right)} \\ &= \frac{(sgn)}{\sqrt{(sgn) g_{\rho\sigma} \frac{d\gamma^\rho}{d\lambda} \frac{d\gamma^\sigma}{d\lambda}}} g_{\alpha\nu} \frac{d\gamma^\nu}{d\lambda}. \end{aligned}$$

Using these, the Hamiltonian function can be sought through the usual Legendre transformation

$$H(\gamma^\alpha, p_\beta) := p_\mu \frac{d\gamma^\mu}{d\lambda}(\gamma^\alpha, p_\beta) - L\left(\gamma^\alpha, \frac{d\gamma^\rho}{d\lambda}(\gamma^\alpha, p_\beta)\right).$$

However, notice that replacing p_μ with its expression in terms of the generalized velocities $d\gamma^\mu/d\lambda$ leads to

$$H = \frac{(sgn) g_{\mu\nu} \frac{d\gamma^\nu}{d\lambda} \frac{d\gamma^\mu}{d\lambda}}{\sqrt{(sgn) g_{\rho\sigma} \frac{d\gamma^\rho}{d\lambda} \frac{d\gamma^\sigma}{d\lambda}}} - \sqrt{(sgn) g_{\mu\nu}(\gamma^\alpha) \frac{d\gamma^\mu}{d\lambda} \frac{d\gamma^\nu}{d\lambda}} = 0.$$

That is, the Hamiltonian function is exactly zero. This is typically the case with reparameterization-invariant (gauge invariant) actions (see Henneaux and Teitelboim [1992]). So it seems that at face value, the Hamiltonian formulation fails. However, there is a well known alternative.

To obtain a well defined Hamiltonian formulation of general-relativistic particle mechanics an alternative Lagrangian formulation is used. Consider the action

$$S_2[\gamma^\mu] = \int_{\lambda_0}^{\lambda_1} d\lambda \ \left(\frac{1}{2} g_{\mu\nu} \frac{d\gamma^\mu}{d\lambda} \frac{d\gamma^\nu}{d\lambda} \right)$$

often called the **energy** of the path γ, because it has the form $\frac{1}{2} m \vec{v} \cdot \vec{v}$ with $m = 1$. This action is not reparameterization invariant, which at first seems like a problem. However, the condition that a path γ_c^μ extremizes this action is

$$\begin{aligned} 0 &= \left. \frac{\delta S_2[\gamma^\nu]}{\delta \gamma^\mu} \right|_{\gamma_c} \\ &= \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left(S_2[\gamma_c^\mu + \epsilon \delta\gamma^\mu] - S_2[\gamma_c^\mu] \right) \\ &= \int_{\lambda_0}^{\lambda_1} d\lambda \ \left(g_{\mu\nu}(\gamma_c) \frac{d\gamma_c^\mu}{d\lambda} \frac{d\delta\gamma^\nu}{d\lambda} + \frac{1}{2} \partial_\rho g_{\mu\nu}(\gamma_c) \frac{d\gamma_c^\mu}{d\lambda} \frac{d\gamma_c^\nu}{d\lambda} \delta\gamma^\rho \right) \end{aligned}$$

which after the usual integration by parts, dropping total derivatives (because of fixed endpoints) and a little re-indexing becomes

$$\int_{\lambda_0}^{\lambda^1} d\lambda \ \left(-g_{\mu\rho} \frac{d^2\gamma_c^\mu}{d\lambda^2} - \partial_\nu g_{\mu\rho} \frac{d\gamma_c^\mu}{d\lambda} \frac{d\gamma_c^\nu}{d\lambda} + \frac{1}{2} \partial_\rho g_{\mu\nu}(\gamma_c) \frac{d\gamma_c^\mu}{d\lambda} \frac{d\gamma_c^\nu}{d\lambda} \delta\gamma^\rho \right) \delta\gamma^\rho = 0.$$

Again, since $\delta\gamma$ is arbitrary, for the integral to vanish exactly, the integrand must vanish. After splitting the second term in the integrand into two pieces, multiplying by the inverse metric, and re-indexing, the integrand becomes

$$\frac{d^2\gamma^\rho}{d\lambda^2} + \frac{1}{2} g^{\rho\sigma} \left(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu} \right) \frac{d\gamma^\mu}{d\lambda} \frac{d\gamma^\nu}{d\lambda} = 0.$$

This is none other than the affinely parameterized geodesic equation. So it turns out that geodesics also extremize the energy integral S_2, so long as they have an affine

parameterization. Therefore if the goal is to obtain such geodesics, the action S_2 is just as good, if not better, than the action S_1.

Furthermore, for this action, the Hamiltonian story has a different ending. Computing the canonically conjugate momentum to the generalized coordinates γ^μ leads to

$$p_\alpha = g_{\alpha\nu}\frac{d\gamma^\nu}{d\lambda}.$$

Then, plugging this expression into the Legendre transform of the Lagrangian results in

$$\begin{aligned} H(\gamma^\alpha, p_\beta) &:= p_\mu \frac{d\gamma^\mu}{d\lambda}(\gamma^\alpha, p_\beta) - L\left(\gamma^\alpha, \frac{d\gamma^\rho}{d\lambda}(\gamma^\alpha, p_\beta)\right) \\ &= p_\mu g^{\mu\nu} p_\nu - \frac{1}{2} g_{\mu\nu} g^{\alpha\mu} p_\alpha g^{\beta\nu} p_\beta \end{aligned}$$

which finally results in

$$\boxed{H(\gamma^\mu, p_\nu) = \frac{1}{2} g^{\alpha\beta}(\gamma^\mu) p_\alpha p_\beta.}$$

This is a rather more confidence inspiring expression. Indeed evaluating Hamilton's equations yields

$$\begin{aligned} \frac{d\gamma^\mu}{d\lambda} &= \frac{\partial H}{\partial p_\mu} \\ &= g^{\mu\nu} p_\nu \\ \frac{dp_\mu}{d\lambda} &= -\frac{\partial H}{\partial \gamma^\mu} \\ &= -\frac{1}{2}\partial_\mu g^{\alpha\beta} p_\alpha p_\beta. \end{aligned}$$

Taking another derivative of the first equation, and plugging both of the equations above gives rise to

$$\frac{d^2\gamma^\mu}{d\lambda^2} = g_{\nu\alpha}\partial_\rho g^{\mu\nu}\frac{d\gamma^\rho}{d\lambda}\frac{d\gamma^\alpha}{d\lambda} - \frac{1}{2} g^{\mu\nu} g_{\alpha\rho} g_{\beta\sigma} \partial_\nu g^{\alpha\beta}\frac{d\gamma^\rho}{d\lambda}\frac{d\gamma^\sigma}{d\lambda}.$$

Using the integration by parts and the fact that $g_{\alpha\beta}g^{\beta\rho} = \delta^\rho_\alpha$, hence its derivative is zero, the equation above becomes

$$\frac{d^2\gamma^\mu}{d\lambda^2} = -\frac{1}{2} g^{\mu\nu}\left(2\partial_\rho g_{\nu\sigma} - \partial_\nu g_{\rho\sigma}\right)\frac{d\gamma^\rho}{d\lambda}\frac{d\gamma^\sigma}{d\lambda}$$

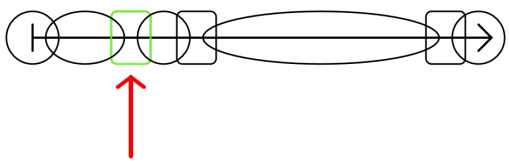
which is simply another way of writing the affinely parameterized geodesic equation. In conclusion, for affinely parameterized geodesics there is a Hamiltonian formulation.

Furthermore, notice that there was no need to specify a (sgn) for the integrand in S_2 to be well defined. This means that there is no need to restrict the paths considered to only those that were always time-like or always space-like, and yet the condition for extremization was again just the geodesics equation, albeit with an affine parameter.

Appendix C: Hierarchical Asymptotic Matching

C.1 Worked Example

In this appendix, I provide the detailed work of developing the intermediate approximation considered in section 9.4, the one used to bridge the global $\epsilon \to 0$ approximation and the local $s \to 1/2$ approximation in the Riemannian sector. In this particular matching attempt, solving the dominant balance equation to obtain even the second term in the approximation becomes exceedingly difficult, so I side-step the exact solutions and go straight to obtaining the asymptotic behavior of this approximation in the two regimes that are used for the matching procedure; namely $r \to \infty$ and $r \to 0$. The results of this work are shown in equations (9.8) and (9.9).



So, starting from the equation

$$S'' + \left(\frac{(n-2)}{s} + \frac{1}{s(1-4s^2)}\right) S' - \left(\frac{1}{\epsilon^2}\frac{(1-4s^2)}{s^2} + \lambda^2(1-4s^2)\right) S = 0$$

and performing the stretching transformation

$$s = \frac{1}{2} - \epsilon^\alpha r, \qquad S(s) = U(r(s))$$

with distinguished limit $\alpha = 2/3$, yields

$$U'' + P(r,\epsilon)U' + Q(r,\epsilon)U = 0 \tag{C.1}$$

where

$$P(r,\epsilon) = -\left(\frac{2(n-2)\epsilon^{2/3}}{1-2r\epsilon^{2/3}} + \frac{1}{2r(1-r\epsilon^{2/3})(1-2r\epsilon^{2/3})}\right)$$
$$Q(r,\epsilon) = -\left(\frac{16r(1-r\epsilon^{2/3})}{(1-2r\epsilon^{2/3})^2} + 4\lambda^2 r\epsilon^2(1-r\epsilon^{2/3})\right).$$

The coefficients $P(r,\epsilon)$ and $Q(r,\epsilon)$ have the following expansions in ϵ, as $\epsilon \to 0$

$$\begin{aligned} P(r,\epsilon) =& -\frac{1}{2r} - \left(2n - \frac{5}{2}\right)\epsilon^{2/3} - \left(4n - \frac{9}{2}\right) r\epsilon^{4/3} + O(\epsilon)^2 \\ Q(r,\epsilon) =& -16r - 48r^2\epsilon^{2/3} - 128r^3\epsilon^{4/3} + O(\epsilon)^2 \end{aligned}$$

The consistent dominant balance is

$$U'' - \frac{1}{2r}U' \sim 16rU$$

whose solution is

$$\exp\left(\pm\frac{8}{3}r^{3/2}\right).$$

This means that U takes the form

$$U = \exp\left(\pm\frac{8}{3}r^{3/2} + B(r)\right)$$

where $B << O(1)$ as $\epsilon \downarrow 0$. Plugging this expression into the equation for U (equation C.1), making sure to include the order $\epsilon^{2/3}$ terms in expansions of the coefficients P and Q, yields

$$B'' + (B')^2 + \left(\pm 8r^{1/2} - \frac{1}{2r} - \left(2n - \frac{5}{2}\right)\epsilon^{2/3}\right) B' = (48r^2 \pm 2(4n-5)r^{1/2})\epsilon^{2/3}$$

The only consistent dominant balance is

$$B'' + \left(\pm 8r^{1/2} - \frac{1}{2r}\right) B' \sim (48r^2 \pm 2(4n-5)r^{1/2})\epsilon^{2/3}, \tag{C.2}$$

but as already discussed in section 9.4, this cannot be solved exactly in terms of analytic function. So instead of solving this equation exactly, I will work first in the $r \to \infty$ limit, and obtain an asymptotic approximation there, and then work in the $r \to 0$ limit an obtain a different asymptotic approximation in that regime. The result will be asymptotic expansions of the same intermediate solution U, which will allow the overall matching procedure to go through (between global $\epsilon \to 0$ approximation and local $s \to 1/2$ approximation).

$r \to \infty$ regime

Starting from the consistent dominant balance for the B term, that is equation C.2, and taking the $r \to \infty$ limit yields

$$B'' \pm 8r^{1/2}B' \sim 48r^2\epsilon^{2/3}$$

There is a consistent two term dominant balance, namely

$$\pm\, 8r^{1/2}B' \sim 48r^2\epsilon^{2/3},$$

whose solution is

$$\pm\, \frac{12}{5}r^{5/2}\epsilon^{2/3}$$

To get the next term, I plug the ansatz

$$U = \exp\left(\pm\frac{8}{3}r^{3/2} \pm \frac{12}{5}r^{5/2}\epsilon^{2/3} + C\right)$$

where $C << B$ as $\epsilon \downarrow 0$ and $r \to \infty$, into the next order U equation, which is

$$U'' - \left(\frac{1}{2r} + \left(2n - \frac{5}{2}\right)\epsilon^{2/3} + \left(4n - \frac{9}{2}\right)r\epsilon^{4/3}\right)U' = (16r + 48r^2\epsilon^{2/3} + 128r^3\epsilon^{4/3})U$$

This results in

$$\begin{aligned} & C''' + (C')^2 + \left(\pm 8r^{1/2} - \frac{1}{2r} + \left(\pm 12r^{3/2} - \left(2n - \frac{5}{2}\right)\right)\epsilon^{2/3} - \left(4n - \frac{9}{2}\right)r\epsilon^{4/3}\right)C' \\ = \;& \pm 8(n-2)r^{1/2}\epsilon^{2/3} + \left(92r^3 \pm (28n - 33)\, r^{3/2}\right)\epsilon^{4/3} \pm 3\,(8n - 9)\, r^{5/2}\epsilon^2 \end{aligned} \tag{C.3}$$

At this point, there is a choice to be made in this algorithm. Following the previous approach of simply dropping higher order terms in ϵ leads to the preliminary dominant balance

$$C'' + \left(\pm 8r^{1/2} - \frac{1}{2r}\right)C' \sim \pm 8(n-2)r^{1/2}\epsilon^{2/3}. \tag{C.4}$$

In the $r \to \infty$ limit this is further reduced to

$$C'' \pm 8r^{1/2}C' \sim \pm 8(n-2)r^{1/2}\epsilon^{2/3} \tag{C.5}$$

whose only consistent dominant balance is

$$\pm\, 8r^{1/2}C' \sim \pm 8(n-2)r^{1/2}\epsilon^{2/3}. \tag{C.6}$$

The solution to this dominant balance is

$$C = (n-2)r\epsilon^{2/3}. \tag{C.7}$$

Notice that this C term is of the same ϵ order as the B term, but it is still much smaller than B in the limit $r \to \infty$. The solution so far is

$$U = \exp\left(\pm\frac{8}{3}r^{3/2} + \left(\pm\frac{12}{5}r^{5/2} + (n-2)r\right)\epsilon^{2/3}\right). \tag{C.8}$$

So what I obtained is really just the second order term in an expansion of what the exact B term would be (had I solved the exact B dominant balance) in the limit of $r \to \infty$. And if I plug this back into to the equation for U, and follow the usual procedure, the next term obtained will also be proportional to $\epsilon^{2/3}$. In fact it is

$$\mp\frac{1}{8}(n-2)r^{-1/2}\epsilon^{2/3}.$$

Continuing in the same fashion leads to a construction of the series expansion of the exact B term. So while I have succeeded in obtaining the asymptotics of the B term, it seems like the terms higher order in ϵ are inaccessible via this method. However, that is not the case. To get to the next order in ϵ, I go back to equation C.3. This time, instead of dropping all higher order terms in ϵ, I first imagine what it would be like if I had a complete solution for B. That is, which terms would such a solution cancel. In this case it would be the first term on the RHS; the term

$$\pm\,8(n-2)r^{1/2}\epsilon^{2/3}. \tag{C.9}$$

The validation for this choice comes from the self-consistency of the next step (as has been done over and over again with the dominant balance technique.) Removing this term, the equation becomes

$$\begin{aligned} & C''' + (C')^2 + \left(\pm 8r^{1/2} - \frac{1}{2r} + \left(\pm 12r^{3/2} - \left(2n - \frac{5}{2}\right)\right)\epsilon^{2/3} - \left(4n - \frac{9}{2}\right)r\epsilon^{4/3}\right)C' \\ = \;& \left(92r^3 \pm (28n-33)\,r^{3/2}\right)\epsilon^{4/3} \pm 3\,(8n-9)\,r^{5/2}\epsilon^2. \end{aligned} \tag{C.10}$$

Now, I proceed as usual. Dropping higher order ϵ terms leads to a preliminary dominant balance

$$C'' + \left(\pm 8r^{1/2} - \frac{1}{2r}\right)C' \sim \left(92r^3 \pm (28n-33)\,r^{3/2}\right)\epsilon^{4/3}$$

and further taking the $r \to \infty$ limit results in

$$C'' \pm 8r^{1/2}C' \sim 92r^3\epsilon^{4/3}.$$

The consistent two-term dominant balance is

$$\pm\, 8r^{1/2}C' \sim 92r^3\epsilon^{3/4}$$

whose solution is

$$C = \pm\frac{23}{7}r^{7/2}\epsilon^{4/3}.$$

Thus the most developed expansion so far, for the intermediate approximation in the limit of $r \to \infty$ is

$$U = \exp\left(\pm\frac{8}{3}r^{3/2} + \left(\pm\frac{12}{5}r^{5/2} + (n-2)r \mp \frac{1}{8}(n-2)r^{-1/2}\right)\epsilon^{2/3} \pm \frac{23}{7}r^{7/2}\epsilon^{4/3}\right). \tag{C.11}$$

A further check that this procedure works comes from the fact that this asymptotic solution to the intermediate approximation problem actually matches the local approximation it was intended to match (see section 9.4).

This procedure can be continued further. To get more terms in the expansion of the B term (the second second term in the $O(\epsilon)$ expansion), simply continue down the first path, doing the usual dominant balance procedure. This will only end when the exact B solution is reached, which only happens if the series terminates after a finite number of terms. To get more terms for the expansion of the C term, once the jump to the next ϵ order has been made by removing the term(s) the exact B term would cancel out, simply continue the usual dominant balance procedure. To move to a higher order ϵ terms, a similar jump must be made again, and so on so forth.

The one very important caveat is that to obtain accurate expansions to a given order, all the terms that affect that order must be included. For example, when working with the order $\epsilon^{4/3}$ term (the C term), it is essential to include the order $\epsilon^{4/3}$ terms in the P and Q expansions. That much is expected. However, what is equally important is that to push the C expansion further, all the B terms that would contribute to the desired r order in C, must be included. That is, for any given ϵ order term (say C, or D, etc.), how far that term can be expanded in orders of r is limited by how far the previous ϵ order term was expanded. So for any given desired accuracy, a hierarchy of terms of more dominant terms are needed, in both orders of ϵ and orders of r.

Without going through the details, since the procedure is very much analogous to what was done above, the $r \to 0$ asymptotics of the intermediate limit equation C.1

is

$$U = \exp\left(\pm \frac{8}{3} r^{3/2} + \left(\pm \frac{4}{5}(4n-5) r^{5/2} - \frac{32}{5}(n-2) r^4 \right) \epsilon^{2/3} \right.$$
$$\left. \pm \frac{1}{7}\left(16n^2 - 24n + 7\right) r^{7/2} \epsilon^{4/3} + O(\epsilon)^2 \right) \qquad \text{(C.12)}$$

Appendix D: Further Calculations

This appendix contains some of the asymptotic analysis calculations that must be performed but are not necessary in the main text. They are referenced in the appropriate sections.

D.1 Lorentzian left intermediate

- The goal here is to obtain the third term in the $r \to 0$ regime. Starting from the U equation, going to order $\epsilon^{4/3}$ in the P and Q expansions, and plugging in

$$U = \exp\left(\pm i\frac{8}{3}r^{3/2} \mp i\frac{12}{5}\epsilon^{2/3}r^{5/2} + C\right)$$

yields

$$C'' + (C')^2 + \left(-\frac{1}{2r} \pm i8r^{1/2} + \frac{3}{2}\epsilon^{2/3} \mp 12i\epsilon^{2/3}r^{3/2} - \frac{7}{2}\epsilon^{4/3}r\right)C'$$
$$\mp 23i\epsilon^{4/3}r^{3/2} + 92\epsilon^{4/3}r^3 \pm 21i\epsilon^2 r^5 = 0.$$

There seems to be no more $\epsilon^{3/2}$ terms, so the B_1 term is exact. The unique dominant balance is

$$C'' - \frac{1}{2r}C' \sim \pm 23i\epsilon^{4/3}r^{3/2}$$

with solution

$$C = \pm\frac{23}{7}i\epsilon^{4/3}r^{7/2} + O[9/2]$$

So, all together, the approximation is

$$U = \exp\left(\pm i\frac{8}{3}r^{3/2} \mp i\frac{12}{5}\epsilon^{2/3}r^{5/2} + \left(\pm\frac{23}{7}ir^{7/2} + O[r^{9/2}]\right)\epsilon^{4/3} + O[\epsilon^2]\right)$$

- Now, I move to obtaining the third term in the $r \to \infty$ regime. Given that the first two solution terms are the same for $r \to 0$ and $r \to \infty$ regimes, I arrive at the same equation for C

$$C'' + (C')^2 + \left(-\frac{1}{2r} \pm i8r^{1/2} + \frac{3}{2}\epsilon^{2/3} \mp 12i\epsilon^{2/3}r^{3/2} - \frac{7}{2}\epsilon^{4/3}r\right)C'$$
$$\mp 23i\epsilon^{4/3}r^{3/2} + 92\epsilon^{4/3}r^3 \pm 21i\epsilon^2 r^5 = 0.$$

The dominant balance is

$$\pm\, 8ir^{1/2}C' \sim -92\epsilon^{4/3}r^3$$

with solution

$$C = \pm\frac{23}{7}i\epsilon^{4/3}r^{7/2} + O[9/2].$$

This is again, the same as the $r \to 0$ regime. So the solution so far is also the same

$$U = \exp\left(\pm i\frac{8}{3}r^{3/2} \mp i\frac{12}{5}\epsilon^{2/3}r^{5/2} + \left(\pm\frac{23}{7}ir^{7/2} + O[r^{9/2}]\right)\epsilon^{4/3} + O[\epsilon^2]\right).$$

It seems like this is saying that there really was no need to do an intermediate matching between $s \to 0$ and $\epsilon \to 0$ solutions. Note that this is not the case when $n \neq 0$

D.2 Lorentzian Right Intermediate

- Continuing where the main text left off, the next step is to plug

$$B = \pm\frac{i}{4}\left(\frac{\sqrt{1+\lambda^2r^2}}{r} - \lambda\sinh^{-1}(\lambda r)\right) - \frac{1}{4}\ln(1+\lambda^2r^2) + C$$

into

$$B'' + (B')^2 + \left(\pm\frac{4i}{\epsilon^2}\sqrt{1+\lambda^2r^2} - \frac{\epsilon^2}{r(4r^2-\epsilon^2)}\right)B'$$
$$+\frac{1}{\epsilon^2}\left(\pm\frac{2i\lambda^2 r}{\sqrt{1+\lambda^2r^2}} - \frac{1+\lambda^2r^2}{r^2}\right) \mp \frac{2i\sqrt{1+\lambda^2r^2}}{r(4r^2-\epsilon^2)} = 0,$$

which yields

$$0 = C'' + (C')^2 + \left(\pm\frac{4i}{\epsilon^2}\sqrt{1+\lambda^2r^2} - \frac{\lambda^2 r}{1+\lambda^2r^2} \mp i\frac{\sqrt{1+\lambda^2r^2}}{2r^2} - \frac{\epsilon^2}{r(4r^2-\epsilon^2)}\right)C'$$
$$-\frac{(1+\lambda^2r^2)}{16r^4} + \frac{5\lambda^4r^2}{4(1+\lambda^2r^2)^2} - \frac{\lambda^2}{2(1+\lambda^2r^2)} \mp \frac{i\epsilon^2}{4}\left(\frac{\sqrt{1+\lambda^2r^2}}{r^3(4r^2-\epsilon^2)}\right) + \frac{\epsilon^2\lambda^2}{2(4r^2-\epsilon^2)(1+\lambda^2r^2)}$$

- The dominant balance is

$$\pm\,\frac{4i}{\epsilon^2}\sqrt{1+\lambda^2r^2}C' \sim \frac{(1+\lambda^2r^2)}{16r^4} - \frac{5\lambda^4r^2}{4(1+\lambda^2r^2)^2} + \frac{\lambda^2}{2(1+\lambda^2r^2)}$$

whose solution yields

$$C = \pm\frac{i\epsilon^2}{192}\left(\frac{(1+\lambda^2r^2)^{3/2}}{r^3} - \frac{4\lambda^2 r}{\sqrt{1+\lambda^2r^2}} - \frac{20\lambda^2 r}{(1+\lambda^2r^2)^{3/2}}\right)$$

D.3 $s \to \infty$ approximation, $n = 2$, $\lambda \neq 0$

- The starting point is the mode equation

$$S'' - \frac{1}{s(4s^2-1)}S' + \left(\frac{1}{\epsilon^2}\frac{(4s^2-1)}{s^2} + \lambda^2(4s^2-1)\right)S = 0$$

- The first step is to perform the typical exponential transformation: $S = \exp(A)$

$$A'' + (A')^2 - \frac{1}{s(4s^2-1)}A' + \left(\frac{1}{\epsilon^2}\frac{(4s^2-1)}{s^2} + \lambda^2(4s^2-1)\right) = 0$$

- The dominant balance is

$$(A')^2 \sim -4\lambda^2 s^2$$

  which yields

$$A = \pm i\lambda s^2 + B$$

  where $B = o\left(s^2\right)$ as $s \to \infty$.

- Plugging this back into equation for A yields

$$B'' + (B')^2 + \left(\pm 4i\lambda s - \frac{1}{s(4s^2-1)}\right)B' \pm 2i\lambda\frac{(4s^2-2)}{(4s^2-1)} + \frac{1}{\epsilon^2}\frac{(4s^2-1)}{s^2} - \lambda^2 = 0.$$

- The dominant balance is

$$\pm\, 4i\lambda s B' \sim -\frac{4}{\epsilon^2} + \lambda^2 \mp 2i\lambda$$

  which yields

$$B = \left(-\frac{1}{2} \pm i\left(\frac{1}{\lambda\epsilon^2} - \frac{\lambda}{4}\right)\right)\ln(s) + C$$

  where $C = o\left(\ln(s)\right)$ as $s \to \infty$.

- Plugging this back into B equation gives

$$\begin{aligned} &C'' + (C')^2 + \left(\pm 4i\lambda s + \frac{2\alpha}{s} - \frac{1}{s(4s^2-1)}\right)C' \\ &+\frac{1}{s^2}\left(\alpha^2 - \alpha - \frac{1}{\epsilon^2}\right) \mp \frac{2i\lambda}{(4s^2-1)} - \frac{\alpha}{s^2(4s^2-1)} = 0 \end{aligned} \tag{D.1}$$

  where

$$\alpha := \left(-\frac{1}{2} \pm i\left(\frac{1}{\lambda\epsilon^2} - \frac{\lambda}{4}\right)\right),$$

- The dominant balance is

$$\pm 4i\lambda s C' \sim \left(-\alpha^2 + \alpha + \frac{1}{\epsilon^2} \pm \frac{i\lambda}{2}\right) \frac{1}{s^2}$$

after plugging in the value of α and simplifying, it becomes

$$C' \sim \left(\frac{1}{2\lambda^2\epsilon^2} \mp i\left(\frac{1}{4\lambda^3\epsilon^4} + \frac{1}{8\epsilon^2\lambda} + \frac{\lambda}{64} - \frac{3}{16\lambda}\right)\right) \frac{1}{s^3}$$

which means

$$C = \left(-\frac{1}{4\lambda^2\epsilon^2} \pm i\left(\frac{1}{8\lambda^3\epsilon^4} + \frac{1}{16\epsilon^2\lambda} + \frac{\lambda}{128} - \frac{3}{32\lambda}\right)\right) \frac{1}{s^2} + D$$

where $D = o(s^{-2})$ as $s \to \infty$.

Bibliography

Arlen Anderson and Bryce DeWitt. Does the topology of space fluctuate? *Foundations of Physics*, 16(2):91–105, 1986. 8

Niels Arley. *A note on the foundations of geometrical optics*. Munksgaard, 1945. 16

Vladimir Igorevich Arnold. *Mathematical methods of classical mechanics*, volume 60. Springer Science & Business Media, 1989. 19, 22

David Atkatz and Heinz Pagels. Origin of the universe as a quantum tunneling event. *Physical Review D*, 25(8):2065, 1982. 6

Carl M Bender and Steven A Orszag. *Advanced mathematical methods for scientists and engineers I*. Springer Science & Business Media, 1999. 41, 44, 48, 74, 82

Nicholas David Birrell and Paul Charles William Davies. *Quantum fields in curved space*. Number 7. Cambridge university press, 1984. 102, 139

G.F. Chew. Analyticity as a fundamental principle in physics. 1 1965. doi: 10.2172/4586375. 53

Sidney Coleman. The uses of instantons. In *The whys of subnuclear physics*, pages 805–941. Springer, 1979. 6

Sidney Coleman and Frank De Luccia. Gravitational effects on and of vacuum decay. *Physical Review D*, 21(12):3305, 1980. 6

Peter Collas. General relativity in two-and three-dimensional space–times. *American Journal of Physics*, 45(9):833–837, 1977. 11

Complete Dictionary of Scientific Biography. Huygens, christiaan (also huyghens, christian), 2008. URL `http://www.encyclopedia.com/doc/1G2-2830902105.html`. 13

Fay Dowker. *Topology change in quantum gravity*. University Press, Cambridge, 2003. 8

Tevian Dray, Corinne A Manogue, and Robin W Tucker. Particle production from signature change. *General Relativity and Gravitation*, 23(8):967–971, 1991. 8

Tevian Dray, Corinne A Manogue, and Robin W Tucker. Boundary conditions for the scalar field in the presence of signature change. *arXiv preprint gr-qc/9501034*, 1995. 8

Tevian Dray, George Ellis, Charles Hellaby, and Corinne A Manogue. Gravity and signature change. *General Relativity and Gravitation*, 29(5):591–597, 1997. 8

S Habib Mazharimousavi, Ashkan Roozbeh, and M Halilsoy. Electromagnetic wave propagation through inhomogeneous material layers. *Journal of Electromagnetic Waves and Applications*, 27(16):2065–2074, 2013. 15

Jonathan J Halliwell. Introductory lectures on quantum cosmology. In *Quantum cosmology and baby universes*, pages 159–243. World Scientific, 1991. 7

James B Hartle and Stephen W Hawking. Wave function of the universe. *Physical Review D*, 28(12):2960, 1983. 6

Stephen W Hawking and Roger Penrose. The singularities of gravitational collapse and cosmology. In *Proc. R. Soc. Lond. A*, volume 314, pages 529–548. The Royal Society, 1970. 6

Stephen William Hawking and Ian L Moss. Supercooled phase transitions in the very early universe. *Physics Letters B*, 110(1):35–38, 1982. 6

Marc Henneaux and Claudio Teitelboim. *Quantization of gauge systems*. Princeton university press, 1992. 151

Mark H Holmes. *Introduction to perturbation methods*, volume 20. Springer Science & Business Media, 2012. 72

Ali Kaya and Merve Tarman. Cosmological backreaction of a quantized massless scalar field. *Journal of Cosmology and Astroparticle Physics*, 2012(01):040, 2012. 139

Tomi S Koivisto and Tomislav Prokopec. Quantum backreaction in evolving flrw spacetimes. *Physical Review D*, 83(4):044015, 2011. 139

Corinne A Manogue, ED Copeland, and Tevian Dray. The trousers problem revisited. *Pramana*, 30(4):279–292, 1988. 8

Viatcheslav Mukhanov and Sergei Winitzki. *Introduction to quantum effects in gravity*. Cambridge University Press, 2007. 102

T Padmanabhan. Semiclassical approximations for gravity and the issue of backreaction. *Classical and Quantum Gravity*, 6(4):533, 1989. 138

Guillem Pérez-Nadal, Albert Roura, and Enric Verdaguer. Backreaction from nonconformal quantum fields in de sitter spacetime. *Classical and Quantum Gravity*, 25(15):154013, 2008. 139

GC Rossi and M Testa. Ground state wave function from euclidean path integral. *Annals of Physics*, 148(1):144–167, 1983. 6

Erwin Schrödinger. An undulatory theory of the mechanics of atoms and molecules. *Physical Review*, 28(6):1049, 1926. 26

Edward P Tryon. Is the universe a vacuum fluctuation? *Nature*, 246(5433):396, 1973. 6

Alexander Vilenkin. Creation of universes from nothing. *Physics Letters B*, 117(1-2): 25–28, 1982. 6

Robert M Wald. The back reaction effect in particle creation in curved spacetime. *Communications in Mathematical Physics*, 54(1):1–19, 1977. 138